普通高等教育“十五”国家级规划教材

高等代数

（第二版）

下　册

丘维声

高等教育出版社

策　　划　李　蕊
编　　辑　李　蕊
封面设计　刘晓翔
责任绘图　杜晓丹
版式设计　王艳红
责任校对　俞声佳
责任印制

内容简介

本书是高等学校的主干基础课‘高等代数’课程的教材．全书分上、下册．上册共六章，内容包括：线性方程组，行列式，数域 K 上 n 维向量空间 K^n，矩阵的运算，欧几里得空间 R^n，矩阵的相抵与相似，二次型与矩阵的合同．下册共四章．内容包括：多项式环，线性空间，线性映射（包括线性变换和线性函数），具有度量的线性空间（包含欧几里得空间，酉空间，正交空间，辛空间）．本书按节配置适量习题，书末附有习题答案与提示．

本书精选了高等代数课程的教学内容，渗透了现代数学研究结构和态射（即保持运算的映射）的观点，深入浅出，简明易懂，并且注重培养学生数学的思维方式，重视知识的应用．

本书可作为综合大学、理工科大学和师范院校的高等代数课程的教材．

图书在版编目(CIP)数据

高等代数．下 丘维声．—2版．—北京：高等教育出版社，2003.7
ISBN 7-04-011877-7

Ⅰ．高...　Ⅱ．丘...　Ⅲ．高等代数-高等学校-教材　Ⅳ．O15

中国版本图书馆 CIP 数据核字(2003)第043488号

出版发行	高等教育出版社	购书热线	010-64054588
社　　址	北京市西城区德外大街4号	免费咨询	800-810-0598
邮政编码	100011	网　　址	http: www.hep.edu.cn
总　　机	010-82028899		http: www.hep.com.cn
经　　销	新华书店北京发行所		
排　　版	高等教育出版社照排中心		
印　　刷			
		版　　次	1996年6月第1版
开　　本	787×960　1 16		年　月第　版
印　　张	14.75	印　　次	年　月第　次印刷
字　　数	270 000	定　　价	17.30元

第二版前言

高等代数课程是大学数学科学学院(或数学系、应用数学系)的主干基础课之一,在一年级上、下两个学期讲授.学生从中学进入大学,有一个学习方法的适应过程.中学阶段学习的数学都是比较具体的对象.我们因势利导,在本套教材中把高等代数研究的具体对象放在前半部分,而把抽象对象放在后半部分.在上册讲述线性代数研究的具体对象:线性方程组,数域 K 上 n 元有序数组的向量空间 K^n,矩阵的运算, K^n 到 K^s 的线性映射,欧几里得空间 R^n,矩阵的相抵关系、相似关系,矩阵的特征值和特征向量,二次型与矩阵的合同关系.在下册讲述多项式环,任意域上的线性空间,线性映射(包括线性变换和线性函数),具有度量的线性空间(包含欧几里得空间,酉空间,以及正交空间和辛空间).我们在近十年的教学实践中,采取上述教学内容体系,使广大学生比较顺利地学到了高等代数的理论和方法,提高了高等代数课的教学质量.

下册一开始(即第 7 章)讲多项式环.我们由浅入深把中学讲的多项式提高到数域 K 上一元多项式的概念上.在讲了一元多项式的加法和乘法的定义以及它们满足的运算法则后,通过比较整数集 Z、数域 K 上所有一元多项式组成的集合 $K[x]$、数域 K 上所有 n 级矩阵组成的集合 $M_n(K)$之间的共同点,抽象出环的概念.接着讲述了数域 K 上一元多项式环 $K[x]$的通用性质:若 R 是一个有单位元 $1'$的交换环, R_1 是 R 的一个子环,且 $1'\in R_1$, K 到 R_1 有一个双射 τ 保持加法与乘法运算,对于任意给定的 $t\in R$,令 $\sigma_t\left(\sum_{i=0}^{n} a_i x^i\right)=\sum_{i=0}^{n}\tau(a_i)t^i$,则 σ_t 是 $K[x]$到 R 的映射,且 σ_t 保持加法与乘法运算,称 σ_t 是 x 用 t 代入. $K[x]$的这一通用性质表明,我们只要把一元多项式环 $K[x]$中有关加法与乘法的等式研究清楚了,通过不定元 x 用环 R 中任一元素 t 代入,就可以得到环 R 中有关加法与乘法的等式.于是我们在本章以研究一元多项式环 $K[x]$的有关加法与乘法的等式(即研究 $K[x]$的结构)为主线.首先讲了整除的概念和性质,讲了带余除法,讲了最大公因式与互素的概念和性质;然后讲了不可约多项式的概念和性质,唯一因式分解定理,重因式的概念和判别,接着分别决定了复数域、实数域上的所有不可约多项式,讲了有理数域上不可约多项式的判别.在讲完一元多项式环 $K[x]$后,我们又讲了多元多项式环,着重讲了对称多项式.本章的最后一节以模 4 剩余类环为例讲了模 m 剩余类环,讲了域的概念,模 p 剩余类域(p 是素数),介绍了域的特征的概念.最后指出,类似于数域 K 上的一元(多

元)多项式,可以定义任一域 F 上的一元(多元)多项式,并且有关数域 K 上一元多项式环 $K[x]$的结论,只要在它的证明中没有用到这个域含有无穷多个元素,那么它对于任一域 F 上的一元多项式环 $F[x]$也成立.此外,还需要注意,如果域 F 的特征为素数 p,则 F 的任一元素的 p 倍都等于零.由上述看出,我们在处理多项式理论这一模块上,渗透了现代数学的观点:研究结构和态射(即保持运算的映射).由于我们在本章最后一节引进了域的概念,介绍了模 p 剩余类域,因此,我们在后面各章中就可以讲任意域上的线性空间及其线性映射的理论,就可以讲任意域上的线性空间中如何引进度量概念.这种讲法是信息时代的要求,因为在信息的可靠与安全问题(例如,纠错编码与密码)中,需要用到有限域的知识,以及有限域上的线性空间的理论.

线性代数是研究线性空间和线性映射的理论.我们在上册讲了具体的线性空间:数域 K 上 n 元有序数组的向量空间 K^n;讲了具体的线性映射: K^n 到 K^s 的线性映射 $A(\alpha)=A\alpha$,其中 A 是数域 K 上 $s\times n$ 矩阵.在下册我们讲抽象的线性空间:任意域 F 上的线性空间;讲抽象的线性映射:域 F 上线性空间 V 到 V' 的线性映射,其中包括域 F 上线性空间 V 上的线性变换和线性函数.学生在大学学习了一个学期后,学习能力有了提高,这样我们就可以在下册讲线性空间时,不停留在线性空间的定义,以及线性相关和线性无关的定义上,而是以研究线性空间的结构为主线.在第 8 章的第 1 节,我们在讲了线性空间的定义和简单性质,以及线性相关和线性无关,极大线性无关组和向量组的秩的定义和性质之后,就着重研究线性空间的结构,指出任一线性空间的结构由它的一个基所决定,而维数对于研究有限维线性空间的结构起着重要作用.在 n 维线性空间 V 中取定一个基后,每一个向量都有它的坐标,且在不同基下的坐标之间有坐标变换公式.在第 2 节我们又利用子空间来刻画线性空间的结构,在讲了子空间的结构,子空间的交与和以及它们的维数公式之后,着重讲子空间的直和,如果 V 的两个(或若干个)子空间的直和等于 V,那么这也刻画了 V 的结构.在第 3 节我们从域 F 上 n 维线性空间 V 与 F^n 有相同的性质,引出线性空间同构的概念,推导出域 F 上两个有限维线性空间同构的充分必要条件是它们的维数相同,于是域 F 上有限维线性空间的同构类与非负整数之间有一个一一对应.由此看出,有限维线性空间的结构是如此简单!我们还讲了在域 F 上的 n 维线性空间 V 中,取定一个基后,向量到它的坐标的对应 σ 是 V 到 F^n 的一个同构映射,并且 V 的任一子空间 U 在 σ 下的象 $\sigma(U)$ 与 U 的维数相同.利用这个结论可以把对于域 F 上任一 n 维线性空间的性质的研究,归结为对于 F^n 的性质的研究.我们通过例题与习题(包括以后几章的有关习题)让学生掌握这一方法.在第 4 节我们讲了商空间的概念及其维数公式.

在第 9 章我们讲线性映射(包括线性变换和线性函数)的理论.首先从几何

空间在 xOy 平面上的正投影 $P:(x,y,z)\longmapsto(x,y,0)$,以及连续函数 $f(x)$的定积分$\int_a^b f(x)dx$ 引出线性映射的概念;接着讲线性映射的运算,指出域 F 上线性空间 V 到 V'的所有线性映射组成的集合成为域 F 上的一个线性空间,而域 F 上线性空间 V 上的所有线性变换组成的集合既是域 F 上的一个线性空间,又是一个有单位元的环,这是从总体上研究线性映射的结构.其次我们又研究单个线性映射的结构.一方面研究由 V 到 V'的一个线性映射 A 决定的两个子空间:A 的核 $\mathrm{Ker}\ A$(它是 V 的一个子空间)和 A 的象 $\mathrm{Im}\ A$(它是 V'的一个子空间),推导出线性映射的维数公式 $\dim V=\dim \mathrm{Ker}\ A+\dim \mathrm{Im}\ A$;另一方面研究域 F 上 n 维线性空间 V 上的线性变换 A 的矩阵表示,以及 V 到 V'的线性映射 A 的矩阵表示,着重研究 V中是否存在一个基,使得线性交换 A在此基下的矩阵具有简单的形式.从线性变换 A 的特征值和特征向量的概念推导出,A 可对角化的充分必要条件是 V 可以分解成 A 的特征子空间的直和.而 A 的特征子空间 V_{λ_i}中每个向量在 A 下的象仍在 V_{λ_i} 中,由此引出 A 的不变子空间的概念,并且指出研究不可对角化的线性变换 A 的结构,其思路是研究 V 能不能分解成 A 的不变子空间的直和.由于对于任意 $f(x)\in F[x]$,都有 $\mathrm{Ker}\ f(A)$是 A 的不变子空间,并且如果 $f(x)$能分解成两两互素的多项式 $f_1(x), f_2(x),\cdots, f_s(x)$的乘积,则有

$$\mathrm{Ker}\ f(A)=\mathrm{Ker}\ f_1(A)\quad \mathrm{Ker}\ f_2(A)\quad \cdots\quad \mathrm{Ker}\ f_s(A),$$

且 $\mathrm{Ker}\,0=V$,因此如果能找到一个多项式 $f(x)$,使得 $f(A)=0$,且 $f(x)$有上述分解,则 $V=\mathrm{Ker}\ f_1(A)\quad\cdots\quad \mathrm{Ker}\ f_s(A)$.于是引出 A 的零化多项式的概念,进而引出 A 的最小多项式的概念.利用 A 的最小多项式 $m(x)$在 $F[x]$中的因式分解,研究了 A 在 V 的适当基下的矩阵的简单形式:如果 $m(x)$在 $F[x]$中能分解成一次因式的乘积,则 V 中存在一个基,使得 A 在此基下的矩阵为 Jordan 形矩阵,称它是 A 的 Jordan 标准形.在证明这一结论时,我们把它归结为研究幂零变换的结构.此外,我们把线性函数也放在第 9 章中,因为域 F 上线性空间 V 上的线性函数 f就是 V 到 F 的一个线性映射,这样我们可以把有关线性映射的结论直接用到线性函数的性质和线性函数空间(V 的对偶空间)的结构的研究上.

在第 10 章我们讨论如何分别在实数域、复数域、任意域上的线性空间中引进度量概念,研究具有度量的线性空间的结构,并且研究保持度量的线性变换的性质.在这一章的第 1 节讲双线性函数的概念和性质,以及对称双线性函数和斜对称双线性函数的结构.第 2 节对于实数域上的线性空间 V,把 V 上的一个正定对称双线性函数称为 V 的一个内积;V 上如果给定了一个内积,则称 V 是一个实内积空间,有限维的实内积空间称为欧几里得空间.在实内积空间 V 中,可

以引进长度、角度、正交、距离等度量概念；在欧几里得空间 V 中，存在标准正交基，它比一般的基有许多优越之处：计算内积简单，向量的坐标的分量可以用内积表达.第 3 节利用有限维子空间 U 及其正交补 $U^{\perp}$ 来刻画实内积空间 V 的结构：$V = U \quad U^{\perp}$，由此引出正交投影的概念，并介绍其应用，第 4 节研究实内积空间中的正交变换和对称变换.第 5 节讨论在复数域上的线性空间中如何引进内积的概念，给定了一个内积的复线性空间称为酉空间.有关实内积空间的许多结论可以平行搬到酉空间上.第 6 节讨论如何在任意域上的线性空间中引进度量概念：域 F 上线性空间 V 如果指定了一个对称双线性函数 f，则称 V 是一个正交空间；如果指定了一个斜对称双线性函数，则称 V 是一个辛空间.这一节加了 * 号，不必在课堂上讲，供有兴趣的学生自己阅读.

我们还写了五个阅读材料，前三个是关于整数环中的带余除法，最大公因数，唯一因子分解定理.阅读材料四是利用有限维线性空间同构的充分必要条件证明有限域的元素个数是一个素数的方幂.阅读材料五是利用线性空间的子空间及其陪集讨论线性码的编码和译码方法.这些阅读材料不必在课堂上讲，供学生自己阅读.

下册的第二版比第一版在内容上作了精选，由原来共八章精选成四章，着重讲述最基本的和应用广泛的内容.对于每一节配备的习题也作了精心挑选，在书末附有习题解答和提示.

我们认为高等代数课程的教学目标，既要让学生掌握这门课程的基础知识和基本方法，又要培养他们具有数学的思维方式.只有按照数学的思维方式去学习数学，才能学好数学；而且学会数学的思维方式，有助于他们把今后肩负的工作做好，从而使他们终身受益.本书按照数学的思维方式编写每一节的内容，使学生在学习高等代数知识的同时，受到数学思维方式的熏陶，日积月累地培养学生的数学思维方式.

本书(上册和下册)可作为综合大学、理工科大学和师范院校的数学科学学院(或数学系、应用数学系等)的高等代数课程的教材.上册供第一学期使用，下册供第二学期使用.每学期的周学时可为 4 + 2 或 4 + 1 或 4(4 + 2 是指每周讲课 4 学时，习题课 2 学时，4 + 1 的含意类似).

作者衷心感谢本书的责任编辑李蕊和胡乃　同志，他们为本书的编辑出版付出了辛勤劳动.

作者热诚欢迎广大读者对本书提出宝贵意见.

丘维声
于北京大学数学科学学院
2003 年 4 月

序　　言

为了把学生培养成为面向 21 世纪的高水平人才,作者积多年讲授高等代数、抽象代数和群表示论等课程的经验以及从事科研工作的体会,写了一套高等代数讲义,用这套讲义给北京大学数学系和概率统计系 94 级学生讲授高等代数课,取得了很好的教学效果.接着又给这两个系的 95 级学生讲授此课,进一步修改这套讲义,现分上、下两册出版.

这套教材从我国的实际情况出发,面向 21 世纪,尝试对高等代数的教学内容进行一些改革,主要有以下几方面:

努力使教材现代化.21 世纪的人才需要掌握现代数学的思想和方法.为此,本书注意渗透现代数学的一些基本思想和观点.例如,用等价关系把集合划分的思想;从代数结构着眼处理问题的思想;同构分类的思想;态射(保持运算的映射)的观点等.用现代的观点组织和讲授传统的教学内容.例如,通过讨论子空间的结构证明线性方程组有解判别定理;按照矩阵的相抵关系、相似关系、合同关系分别讨论矩阵的相抵分类、相似分类和合同分类,并且寻求每一种关系下的完全不变量;运用线性空间的同构分类思想证明域 F 上任一 n 维线性空间 V 与它的对偶空间 V^* 同构,以及 V 与它的双重对偶空间 V^{**} 同构;在讲一元多项式的概念时,用态射的观点阐述一元多项式环的通用性质;用环同构的观点讨论数域 K 上的多项式与多项式函数之间的关系等等.本书还注意渗透现代数学的一些基本概念.例如,线性流形、商集等概念;结合高等代数的具体对象水到渠成地先后引进了抽象代数的一些基本概念:在一元多项式的概念之后引进环的概念;在讲完多项式环之后引进任意域和有限域的概念,以及域的特征的概念;在讲了线性变换的运算后引进域上的代数的概念;在最后一章当学生已经熟悉了正交变换、酉变换和辛变换的性质后,引进群和子群的概念.

力图在教材中体现代数与几何、分析的联系.21 世纪的数学,分析、代数、几何将会更加相互渗透和有机结合.因此要使学生从大学一年级开始就逐步培养把代数与几何、分析联系起来的能力.书中注意从几何直观或分析背景引出高等代数讨论的问题;在讲述高等代数的概念时列举几何或分析的例子;把高等代数的结论应用于解决几何或分析的问题.例如,介绍了行列式的几何意义;从几何空间的结构引出向量空间的基的概念;运用线性方程组的理论解决一些几何问题;从平面旋转的合成引出矩阵乘法的定义;从二次曲面方程的化简引出实对称矩阵的对角化以及实二次型通过正交替换化成标准形的问题,并且运用所得到的代数结论解决二次曲面方程的化简问题;从函数极值问题引出正定(负定)二

次型的概念,并且运用正定(负定)矩阵解决多元函数的极值问题;在讲线性相关性时,讲述了 n 个 n - 1 次可微函数线性无关的充分条件;从几何空间中的例子引出商空间的概念;等等.为了将线性代数的理论应用到分析上,为泛函分析打下基础,本书讨论的线性空间可以是无限维的,尽量不加有限维的限制.

注重联系实际,加强应用.面向 21 世纪,数学系不仅要培养从事数学科研和教学的人才,而且要培养在其他领域工作的人才.因此要努力培养学生运用数学理论解决实际问题和其他领域中的问题的能力.本书在讲完线性方程组的理论后,用它解决平板受热问题;讲了矩阵可对角化的条件后,用它解决色盲遗传问题;在讲了矩阵的运算之后,解决区组设计、图论、数论中的一些问题;在讲了一元多项式环的通用性质后,用它证明组合数的一些公式等等.考虑到矩阵在实际问题和许多领域中有广泛应用,本书不仅把矩阵贯穿始终,而且把矩阵的运算,矩阵的相抵分类、相似分类、合同分类集中在一起讲授,并且加强了矩阵的分块,矩阵的"打洞"以及巧用特殊矩阵的训练;讲述了 Binet - Cauchy 公式及其应用,考虑到有限域上的线性空间在计算机以及通讯编码中有重要应用,书中讨论的线性空间是任意域上的,不局限于数域.考虑到现代物理以及一些数学分支中的需要,加强了酉空间的内容,介绍了作为爱因斯坦相对论基础的 Minkowski 空间,并且讨论了一般的正交空间,以及辛空间.

提高数学素质,加强能力培养.21 世纪所需要的人才应当有较高的数学素质和较强的分析问题、解决问题的能力.数学素质包括提出数学问题、理解力、逻辑思维、抽象思维、创造性等几个方面.为了从大学一年级开始就着力培养学生的数学素质,本书在每一单元的开头都要提出问题,然后阐述解决问题的想法,经过抽象思维和逻辑思维一步一步地去解决这些问题.书中特别注意讲清楚想法(idea).例如,在研究线性方程组有无解的判定时为什么会想到去研究 n 元有序数组的向量空间的结构?在讨论有理系数多项式的因式分解时怎么会想到引进本原多项式的概念?在研究不可以对角化的线性变换的结构时如何想到最小多项式的因式分解并且讨论相应线性变换的多项式核之间的关系?复线性空间上内积的定义为什么与实线性空间上内积的定义不同?在任意域上的线性空间中如何引进度量?为什么只有两种度量:对称双线性函数或者斜对称双线性函数,而不用一般的双线性函数?等等,书中都给予了清晰的回答.为了培养学生的阅读、理解能力和扩大知识面,书中有较多的例题,并且有密切配合正文的加"*"号的章节内容和一些阅读材料.有些例题不必在课堂上讲授,留给学生自己看.加"*"号的内容和阅读材料不作为教学要求,供有兴趣的学生自学.本书配备了相当丰富的习题(每一节后面配有习题,每一章后面还有补充题),有的是为了使学生理解正文的概念,掌握正文中的定理和方法,学会重要的解题方法和技巧;有的是正文内容的补充和拓宽;有的是应用.通过做题可以培养学生分析问

题和解决问题的能力.习题中加“ * ”号的题以及补充题不作为教学要求,供学生选做.

本书内容的安排力求符合人们认识事物的客观规律.学生从中学进入大学,有一个学习方法的适应过程.为了帮助学生树立信心适应大学的学习,我们把高等代数研究的具体对象放在前半部分,而把抽象对象放在后半部分.全书分为四部分.第一部分是线性方程组,n 元有序数组的向量空间和矩阵的理论,包括线性方程组的解法、行列式、n 元有序数组的向量空间、线性方程组的理论、矩阵的运算、矩阵的相抵分类与相似分类、二次型与矩阵的合同分类.第二部分是一元多项式环与多元多项式环的理论.第三部分是线性空间和线性映射的理论,包括任意域上的线性空间、线性映射和线性变换、线性变换的 Jordan 标准形、线性函数和双线性函数.第四部分是具有度量的线性空间的理论,包括欧几里得空间、酉空间、正交空间、辛空间.

本书可作为大学数学系、概率统计系、应用数学系的高等代数教材,上、下册共讲授两个学期;还可作为大专院校有关教师和学生的参考书.

作者衷心感谢刘旭峰博士,他通读了全书,并提出了一些宝贵的建议.

特别要感谢本书的责任编辑胡乃　,他为本书的编辑出版付出了辛勤的劳动.

书中可能会有考虑不周和疏漏之处,热诚欢迎同行和读者批评指正.

丘维声

1996 年 2 月于北京大学燕北园

致谢:

本书上、下册经国家教委评审组和国家教委高等学校数学与力学教学指导委员会基础数学指导组评议和审定为国家级重点教材,作者特此向他们表示衷心感谢.

丘维声

1997 年 11 月

目　录

第7章　多 项 式 环

在古典代数学里,多项式的求根是一个中心问题.在近世代数学里,多项式环是一类重要的环.在数学分析中,用多项式函数逼近一般的 n 阶可微函数.在当今信息时代,多项式在计算机科学、现代通信、编码和密码等领域中都有重要应用.

本章研究一元多项式环和多元多项式环的结构.最后一节将介绍有限域的概念.

§1　一元多项式环

观察下列表达式有什么不同之处:

$$-x^3+x^2+\frac{1}{2},\quad \text{其中 } x \text{ 是一个符号;} \tag{1}$$

$$-i^3+i^2+\frac{1}{2},\quad \text{其中 } i = \quad -1; \tag{2}$$

$$-A^3+A^2+\frac{1}{2}I, \text{其中 } A=\begin{matrix}1 & 0\\ 0 & -1\end{matrix}. \tag{3}$$

经过简单的计算,从(2)、(3)式分别得出

$$-i^3+i^2+\frac{1}{2}=i-\frac{1}{2}, \tag{4}$$

$$-A^3+A^2+\frac{1}{2}I=-A+\frac{3}{2}I. \tag{5}$$

由此看出,对于 i 的表达式或矩阵 A 的表达式来说,相等的两个表达式可以含有不同的项.而对于 x 的表达式,为了使它有广泛应用,希望它不出现这种情况,为此数学上引出如下概念:

定义1　设 K 是一个数域, x 是一个不属于 K 的符号.任意给定一个非负整数 n,在 K 中任意取定 $a_0, a_1, \cdots, a_n$,表达式

$$a_n x^n + a_{n-1}x^{n-1}+\cdots+a_1x+a_0 \tag{6}$$

称为数域 K 上的一个一元多项式,如果它具有下述性质:两个这种形式的表达式相等当且仅当它们除去系数为零的项外含有完全相同的项,而系数为零的项允许任意删去和添进来.此时,符号 x 称为不定元.

系数全为零的多项式称为零多项式,记为 0.

在多项式(6)中, $a_i x^i$ 称为 i 次项, $i = 1, \cdots, n$; a_0 称为零次项, 也称为常数项.

从定义 1 知道,数域 K 上两个一元多项式相等当且仅当它们的同次项的系数都相等.

我们常常用 $f(x), g(x), \cdots$ 或 $f, g, \cdots$ 等表示一元多项式.

设 $f(x)$ 表示多项式(6).如果 $a_n \neq 0$,则称 $a_n x^n$ 为多项式 $f(x)$ 的首项; n 称为 $f(x)$ 的次数,记作 $\deg f(x)$ 或 $\deg f$.

零多项式的次数定义为 $-\infty$,并且规定:

$$(-\infty) + (-\infty) = -\infty,$$

$$(-\infty) + n = -\infty,$$

$$-\infty < n,$$

其中 n 是任意非负整数.

零次多项式是 a,其中 $a \in K$ 且 $a \neq 0$.

我们把数域 K 上的所有一元多项式组成的集合记作 $K[x]$.在 $K[x]$ 中规定加法与乘法运算如下:

设

$$f(x) = \sum_{i=0}^{n} a_i x^i, \quad g(x) = \sum_{i=0}^{m} b_i x^i,$$

不妨设 $n \geq m$,则

$$f(x) + g(x) \overset{\text{def}}{=} \sum_{i=0}^{n} (a_i + b_i) x^i, \tag{7}$$

$$f(x) g(x) \overset{\text{def}}{=} \sum_{s=0}^{n+m} \left(\sum_{i+j=s} a_i b_j \right) x^s, \tag{8}$$

称 $f(x) + g(x)$ 是 $f(x)$ 与 $g(x)$ 的和,称 $f(x) g(x)$ 是 $f(x)$ 与 $g(x)$ 的积.

容易验证上面所定义的多项式的加法与乘法满足下列运算法则: $\forall f, g, h \in K[x]$,有

1° 加法交换律,即 $f + g = g + f$;

2° 加法结合律,即 $(f + g) + h = f + (g + h)$;

3° 零多项式具有性质: $0 + f = f + 0 = f$;

4° 设 $f(x) = \sum_{i=0}^{n} a_i x^i$,定义 $-f(x) = \sum_{i=0}^{n} (-a_i) x^i$,则

$$f + (-f) = (-f) + f = 0,$$

称 $-f$ 是 f 的负元素;

5° 乘法交换律,即 $fg = gf$;

6° 乘法结合律,即$(fg)h=f(gh)$;

7° 零次多项式1具有性质:$1f=f1=f$;

8° 乘法对于加法的分配律:

$$f(g+h)=fg+fh,$$
$$(g+h)f=gf+hf.$$

多项式的减法定义如下:

$$f-g \overset{\text{def}}{=} f+(-g). \tag{9}$$

命题1 设$f(x),g(x)\in K[x]$,则

$$\deg(f\pm g)\leq \max\{\deg f,\deg g\}; \tag{10}$$
$$\deg(fg)=\deg f+\deg g. \tag{11}$$

证明 如果$f=0$或$g=0$,则(10),(11)式显然成立.下面设$f\neq 0$且$g\neq 0$.设

$$f(x)=\sum_{i=0}^{n} a_i x^i,\quad g(x)=\sum_{i=0}^{m} b_i x^i,$$

其中$a_n\neq 0,b_m\neq 0$.于是$\deg f=n,\deg g=m$,不妨设$n\geq m$.由于

$$f(x)\pm g(x)=\sum_{i=0}^{n}(a_i\pm b_i)x^i.$$

因此 $$\deg(f\pm g)\leq n=\max\{\deg f,\deg g\}.$$

由于$a_nb_m\neq 0$,因此$a_nb_mx^{n+m}$是$f(x)g(x)$的首项.从而

$$\deg(fg)=n+m=\deg f+\deg g.$$ x

从命题(1)的证明过程中可看出,如果$f\neq 0$且$g\neq 0$,则$fg\neq 0$.由此得出,多项式的乘法适合

9° 消去律,即如果$fg=fh$且$f\neq 0$,则$g=h$.

证明 从$fg=fh$得,$f(g-h)=0$.由于$f\neq 0$,因此$g-h=0$.即$g=h$. x

观察下列集合有什么共同点:

整数集Z;

数域K上所有一元多项式组成的集合$K[x]$;

数域K上所有n级矩阵组成的集合$M_n(K)$.

在集合Z,$K[x]$,$M_n(K)$中都有加法和乘法运算,并且它们满足加法结合律,加法交换律,有零元素,每个元素有负元素,还满足乘法结合律,乘法对于加法的左、右分配律,有单位元素.

多项式的乘法与矩阵的乘法是不同的.我们需要搞清楚什么是集合的一种运算.

定义2 设S是一个非空集合,从$S\times S$到S的一个映射称为S的一个代

数运算.

从 $\mathbf{Z}$, $K[x]$, $M_n(K)$的共同点,我们抽象出下述概念:

定义 3　设 R 是一个非空集合,如果它有两个代数运算,一个叫做加法,记作 $a+b$,另一个叫做乘法,记作 ab;并且这两个运算满足下列 6 条运算法则: $\forall\, a, b, c\in R$,有

1°　加法结合律,即$(a+b)+c=a+(b+c)$;

2°　加法交换律,即 $a+b=b+a$;

3°　在 R 中有元素 0,使得 $a+0=a$,称 0 是 R 的零元素;

4°　对于 a,在 R 中有元素 d,使得 $a+d=0$,称 d 是 a 的负元素,记作 $-a$;

5°　乘法结合律,即$(ab)c=a(bc)$;

6°　乘法对于加法的左、右分配律,即

$$a(b+c)=ab+ac,$$
$$(b+c)a=ba+ca,$$

则称 R 是一个环.

容易证明,环 R 中的零元素是唯一的;每个元素 a 的负元素是唯一的.

$\mathbf{Z}$, $K[x]$, $M_n(K)$都满足定义 3 的条件,因此它们都是环,分别叫做整数环,数域 K 上一元多项式环,数域 K 上 n 级全矩阵环.

任意一个数域 K 也是环.

环 R 中的减法定义成:

$$a-b \overset{\text{def}}{=} a+(-b). \tag{12}$$

环 R 中的乘法如果还满足交换律,即

$$ab=ba,\ \forall\, a, b\in R,$$

则称 R 是交换环.

环 R 中如果有一个元素 e 具有性质:

$$ea=ae=a,\ \forall\, a\in R, \tag{13}$$

则称 e是 R 的单位元(素).此时称 R 是有单位元的环,在有单位元的环 R 中,单位元素是唯一的,通常就记作 1.

环 R 中的元素 a 称为一个左零因子(右零因子),如果 R 中有元素 $b\neq 0$,使得 $ab=0$($ba=0$).左零因子和右零因子都简称为零因子.容易证明,0 既是左零因子,又是右零因子.称 0 是平凡零因子;其余的零因子称为非平凡的零因子.

如果环 R 没有非平凡的零因子,则 R 称为无零因子环.

有单位元 1($\neq 0$)的无零因子的交换环称为整环.

$\mathbf{Z}$, $K[x]$, K 都是整环. $M_n(K)$不是整环.

在整数环 Z 中,全体偶数组成的集合对于整数的加法和乘法也成为一个环,一般地,我们有下述概念:

定义 4 如果环 R 的一个非空子集 R_1 对于 R 的加法和乘法也成为一个环,则 R_1 称为 R 的一个子环.

从定义 4 知道,如果 R_1 是环 R 的一个子环,则 R 的加法(乘法)也是 R_1 的加法(乘法).再从定义 2 知道,这意味着:

$$a, b\in R_1 \qquad a+b\in R_1, \text{且 } ab\in R_1.$$

此时称 R_1 对于 R 的加法和乘法封闭.

可以证明:

命题 2[①] 环 R 的一个非空子集 R_1 为一个子环的充分必要条件是,R_1 对于 R 的减法与乘法都封闭,即

$$a, b\in R_1 \qquad a-b\in R_1, ab\in R_1.$$ x

$K[x]$中所有零次多项式添上零多项式组成的集合 S,对于多项式的减法与乘法封闭,因此 S 是 $K[x]$的一个子环.显然 $K[x]$的单位元 1 属于 S,从而 1 也是 S 的单位元.我们可以建立数域 K 到 S 的一个对应法则:让非零数 a 对应到零次多项式 a,让数 0 对应到零多项式.显然 σ 是双射,并且 σ 保持加法与乘法运算,即

$$\sigma(a+b)=\sigma(a)+\sigma(b), \quad \forall a, b\in K;$$
$$\sigma(ab)=\sigma(a)\sigma(b), \quad \forall a, b\in K.$$

给定 $A\in M_n(K)$,形如下述的表达式称为数域 K 上矩阵 A 的多项式:

$$a_m A^m + a_{m-1} A^{m-1} + \cdots + a_1 A + a_0 I,$$

其中 m 是非负整数,$a_i\in K$, $i=0,1,\cdots,m$.把数域 K 上矩阵 A 的所有多项式组成的集合记作 $K[A]$,即

$$K[A] \overset{\text{def}}{=} \{a_m A^m+\cdots+a_1 A+a_0 I \mid m\in \mathbf{N}, a_i\in K, i=0,1,\cdots,m\}$$

设 $f(A)=\sum_{i=0}^{m} a_i A^i$, $g(A)=\sum_{i=0}^{l} b_i A^i$.不妨设 $m\geq l$.从矩阵的运算法则可得出

$$f(A)-g(A)=\sum_{i=0}^{m}(a_i-b_i)A^i, \tag{14}$$

$$f(A)g(A)=\sum_{s=0}^{m+l}\left(\sum_{i+j=s} a_i b_j\right)A^s. \tag{15}$$

因此 $K[A]$是 $M_n(K)$的一个子环.显然 $I\in K[A]$.从(15)式容易看出,$f(A)g(A)=g(A)f(A)$.因此 $K[A]$是有单位元的交换环.

---

① 证明参看丘维声编著《高等代数(下册)》(高等教育出版社 1996 年 12 月第 1 版)第 8 页.

$K[A]$中所有数量矩阵组成的集合 W,对于矩阵的减法与乘法封闭,因此 W 是 $K[A]$的一个子环.显然 $I\in W$.我们可以建立数域 K 到 W 的一个对应法则 $\tau: a\longmapsto aI$.显然 τ是双射,并且 τ保持加法与乘法运算;

$$\tau(a+b)=(a+b)I=aI+bI=\tau(a)+\tau(b),$$

$$\tau(ab)=(ab)I=(aI)(bI)=\tau(a)\tau(b).$$

$K[A]$中的等式(14)和(15)与一元多项式环 $K[x]$中减法和乘法的定义相像.这不是偶然的.这说明 $K[A]$与一元多项式环 $K[x]$有联系.这种联系抽象出来就是下面所说的一元多项式环 $K[x]$的通用性质:

定理 3　设 K 是一个数域, R 是一个有单位元 $1'$的交换环, R_1 是 R 的一个子环,且 $1'\in R_1$. K到 R_1 有一个双射 τ,并且 τ保持加法与乘法运算(此时有 $\tau(1)=1'$).任意给定 $t\in R$,令

$$\sigma_t:\quad K[x]\qquad\qquad R$$

$$f(x)=\sum_{i=0}^{n}a_ix^i\longmapsto\sum_{i=0}^{n}\tau(a_i)t^i\overset{\text{def}}{=}f(t),$$

则 σ_t 是 $K[x]$到 R的映射;并且 σ_t 保持加法与乘法运算,即如果

$$f(x)+g(x)=h(x),\qquad f(x)g(x)=p(x),\tag{16}$$

那么有

$$f(t)+g(t)=h(t),\quad f(t)g(t)=p(t);\tag{17}$$

此外,$\sigma_t(x)=t$.我们把映射 σ_t 叫做 x用 t代入.

证明　由于 $K[x]$中每个元素 $f(x)$写成 $\sum_{i=0}^{n}a_ix^i$ 的表示法唯一(除了系数为 0 的项以外),并且 τ是 K 到 R_1 的双射,因此 σ_t 是 $K[x]$到 R 的一个映射.

据 σ_t 的定义,得

$$\sigma_t(x)=\sigma_t(1x)=\tau(1)t=1't=t.$$

设　$f(x)=\sum_{i=0}^{n}a_ix^i$, $g(x)=\sum_{i=0}^{m}b_ix^i$,不妨设 $n\geq m$.

则

$$h(x)=f(x)+g(x)=\sum_{i=0}^{n}(a_i+b_i)x^i,$$

$$p(x)=f(x)g(x)=\sum_{s=0}^{n+m}\Big(\sum_{i+j=s}a_ib_j\Big)x^s.$$

据 σ_t 的定义,得

$$h(t)=\sum_{i=0}^{n}\tau(a_i+b_i)t^i=\sum_{i=0}^{n}[\tau(a_i)+\tau(b_i)]t^i$$

$$=\sum_{i=0}^{n}\tau(a_i)t^i+\sum_{i=0}^{n}\tau(b_i)t^i$$

$$= f(t) + g(t);$$

$$p(t) = \sum_{s=0}^{n+m} \tau\left(\sum_{i+j=s} a_i b_j\right) t^s = \sum_{s=0}^{n+m} \left(\sum_{i+j=s} \tau(a_i)\tau(b_j)\right) t^s;$$

$$f(t)g(t) = \left(\sum_{i=0}^{n} \tau(a_i)t^i\right)\left(\sum_{j=0}^{m} \tau(b_j)t^j\right)$$

$$= \sum_{i=0}^{n}\sum_{j=0}^{m} \tau(a_i)\tau(b_j)t^{i+j} = \sum_{s=0}^{n+m}\left(\sum_{i+j=s} \tau(a_i)\tau(b_j)\right) t^s$$

$$= p(t).$$

因此 σ_t 保持加法与乘法运算. x

定理 3 告诉我们,如果 R 是有单位元的交换环,且 R 有一个子环 R_1 满足定理 3 的条件(这时我们称 R 可看成是 K 的一个扩环),那么一元多项式环 $K[x]$ 中所有通过加法与乘法表示的关系,在不定元 x 用 R 的任一元素 t 代入后仍然保持.因此我们只要把一元多项式环 $K[x]$ 中有关加法与乘法的等式研究清楚了,通过不定元 x 用环 R 中任一元素 t 代入,就可以得到环 R 中有关加法与乘法的等式.这就是一元多项式环 $K[x]$ 的通用性质的含义.

从前面的讨论知道, $K[x]$, $K[A]$(其中 A 是 K 上任一 n 级矩阵)都可以作为定理 3 中的环 R.因此,不定元 x 可以用 x 的任一多项式代入,也可以用矩阵 A 的任一多项式代入,从 $K[x]$ 中已知的有关加法和乘法的等式,得到 $K[x]$ 中又一些有关加法和乘法的等式,或者得到 $K[A]$ 中有关加法和乘法的等式.

例 1 设 B 是数域 K 上 n 级幂零矩阵,其幂零指数为 l. 令 $A = I + kB$, $k \in K$. 证明 A 可逆,并且求 A^{-1}.

证明 在 $K[x]$ 中直接计算可得

$$(1 - x)(1 + x + x^2 + \cdots + x^{l-1}) = 1 - x^l. \tag{18}$$

不定元 x 用 $-kB$ 代入,从(18)式得

$$(I + kB)[I - kB + (-kB)^2 + \cdots + (-kB)^{l-1}] = I - (-kB)^l.$$

由于 $B^l = 0$,因此从上式得

$$(I + kB)(I - kB + k^2B^2 + \cdots + (-1)^{l-1}k^{l-1}B^{l-1}) = I. \tag{19}$$

(19)式表明, $I + kB$ 可逆,并且

$$(I + kB)^{-1} = I - kB + k^2B^2 + \cdots + (-1)^{l-1}k^{l-1}B^{l-1}.$$ x

习 题 7.1

1. 在 $K[x]$ 中,如果 $f(x) = cg(x)$,其中 $c \in K$ 且 $c \neq 0$,试问: $f(x)$ 与 $g(x)$ 的次数有什

么关系?

2. 在 $K[x]$中,如果 $f(x)g(x)=c$,其中 $c\in K$ 且 $c\neq 0$,试问:$f(x)$与 $g(x)$的次数各是多少?

3. 在 $K[x]$中,如果 $f(x)$与 $g(x)$的次数都是3,试问:$f(x)+g(x)$的次数一定是3吗?

4. 设 R 是一个有单位元 $1(\neq 0)$的环,对于 $a\in R$,如果存在 $b\in R$,使得

$$ab=ba=1,$$

则称 a 为可逆元(或称 a 为单位,注意不要与单位元1混淆),称 b 是 a 的逆,记作 a^{-1}.

证明:$K[x]$中一个元素 $f(x)$是可逆元当且仅当 $f(x)$是零次多项式.

*5. 设 R 是有单位元 $1(\neq 0)$的环,证明 R 中的可逆元不可能是零因子.

6. 设

$$A=\begin{pmatrix} 1 & b & b^2 & \cdots & b^{n-2} & b^{n-1} \\ 0 & 1 & b & \cdots & b^{n-3} & b^{n-2} \\ \cdots & \cdots & \cdots & & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & b \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

其中 $b\in K$,求 A^{-1}.

7. 设 $A\in M_n(K)$,并且设 A 的特征多项式为

$$|\lambda I-A|=(\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\cdots(\lambda-\lambda_s)^{l_s},$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是两两不同的复数;$l_1+l_2+\cdots+l_s=n$.证明:对于 K 中任一非零数 k,矩阵 kA 的特征多项式为

$$|\lambda I-kA|=(\lambda-k\lambda_1)^{l_1}(\lambda-k\lambda_2)^{l_2}\cdots(\lambda-k\lambda_s)^{l_s}.$$

由此得出,如果 λ_i 是 A 的 l_i 重特征值,则 $k\lambda_i$ 是 kA 的 l_i 重特征值.

*8. 设 A 和 A 的特征多项式同第7题、证明:A^2 的特征多项式为

$$|\lambda I-A^2|=(\lambda-\lambda_1^2)^{l_1}(\lambda-\lambda_2^2)^{l_2}\cdots(\lambda-\lambda_s^2)^{l_s}.$$

由此得出,如果 λ_i 是 A 的 l_i 重特征值,则 λ_i^2 是 A^2 的 l_i 重特征值.

§2　整除性,带余除法

从一元多项式环的通用性质看到,我们应当尽可能多地得到 $K[x]$中有关加法和乘法的等式,为此需要研究一元多项式环 $K[x]$的结构.从本节开始我们将主要研究 $K[x]$的结构,其中 K 是任一数域.

观察 $K[x]$中两个多项式 $f(x)$与 $g(x)$之间有什么关系:

$$f(x)=x^2-1,\quad g(x)=x-1.$$

显然,

$$f(x)=(x+1)g(x).$$

由此我们抽象出下述概念:

定义1　设 $f(x),g(x)\in K[x]$,如果存在 $h(x)\in K[x]$,使得　$f(x)=$

$h(x)g(x)$，则称 $g(x)$ 整除 $f(x)$，记作 $g(x)\mid f(x)$；否则，称 $g(x)$ 不能整除 $f(x)$，记作 $g(x)\nmid f(x)$.

当 $g(x)$ 整除 $f(x)$ 时，$g(x)$ 称为 $f(x)$ 的因式，$f(x)$ 称为 $g(x)$ 的倍式.

容易看出下列事实：

1°　$0\mid f(x)\qquad\qquad f(x)=0$；

2°　$f(x)\mid 0$，$\forall\, f(x)\in K[x]$；

3°　$b\mid f(x)$，$\forall\, b\in K$ 且 $b\neq 0$，$\forall\, f(x)\in K[x]$.

今后我们把 K 中所有非零数组成的集合记作 K^*.

$K[x]$ 中的整除关系具有反身性（即，$f(x)\mid f(x)$，$\forall\, f(x)\in K[x]$），传递性（即，如果 $f(x)\mid g(x)$ 且 $g(x)\mid h(x)$，则 $f(x)\mid h(x)$），但是不具有对称性（即，从 $f(x)\mid g(x)$ 推不出 $g(x)\mid f(x)$）. 传递性的证明留作习题.

定义 2　在 $K[x]$ 中，如果 $f(x)\mid g(x)$ 且 $g(x)\mid f(x)$，则称 $f(x)$ 与 $g(x)$ 相伴，记作 $f(x)\sim g(x)$.

命题 1　在 $K[x]$ 中，$f(x)\sim g(x)$ 当且仅当存在 $c\in K^*$，使得

$$f(x)=cg(x).$$

证明　充分性. 设 $f(x)=cg(x)$，其中 $c\in K^*$. 则 $g(x)\mid f(x)$. 由于 $g(x)=\dfrac{1}{c}f(x)$，因此 $f(x)\mid g(x)$，从而 $f(x)\sim g(x)$.

必要性. 设 $f(x)\sim g(x)$. 则 $f(x)\mid g(x)$ 且 $g(x)\mid f(x)$. 从而存在 $h_1(x)$，$h_2(x)\in K[x]$，使得

$$g(x)=h_1(x)f(x),\qquad f(x)=h_2(x)g(x).$$

于是

$$f(x)=h_2(x)h_1(x)f(x). \tag{1}$$

如果 $f(x)=0$，则 $g(x)=0$. 从而 $f(x)\sim g(x)$. 下面设 $f(x)\neq 0$. 运用消去律，从(1)式得

$$1=h_2(x)h_1(x). \tag{2}$$

从(2)式得，

$$\deg h_2(x)+\deg h_1(x)=0.$$

因此，$\deg h_1(x)=\deg h_2(x)=0$，从而 $h_2(x)=c$，$c\in K^*$. 于是 $f(x)=cg(x)$.

命题 2　在 $K[x]$ 中，如果 $g(x)\mid f_i(x)$，$i=1,2,\cdots,s$，则对于任意 $u_i(x)\in K[x]$，$i=1,2,\cdots,s$，有

$$g(x)\mid(u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots+u_s(x)f_s(x)).$$

命题 2 的证明留作习题.

$K[x]$ 中，如果 $g(x)$ 不能整除 $f(x)$，那么能有什么样的结论呢？例如，设 $f(x)=x^2$，$g(x)=x-1$. 则

$$f(x)=x^2-1+1=(x+1)g(x)+1.$$

由此受到启发,猜想有下述结论:

定理 3(带余除法)　对于 $K[x]$ 中任意两个多项式 $f(x)$ 与 $g(x)$,其中 $g(x)\neq 0$,在 $K[x]$ 中存在唯一的一对多项式 $h(x), r(x)$,使得

$$f(x) = h(x)g(x) + r(x),\quad \deg r(x) < \deg g(x). \tag{3}$$

证明　存在性.设 $\deg g(x) = m\in N$.

情形 1. $m = 0$.此时 $g(x) = b, b\in K^*$.于是

$$f(x) = \frac{1}{b}f(x)\ b + 0,\quad \deg 0 < \deg g(x),$$

情形 2. $m > 0$,且 $\deg f(x) < m$.则

$$f(x) = 0\ g(x) + f(x),\quad \deg f(x) < \deg g(x).$$

情形 3. $m > 0$,且 $\deg f(x)\geq m$.对被除式的次数 n 作数学归纳法.假设对于次数小于 n 的被除式,命题的存在性部分成立.现在来看 n 次多项式 $f(x)$.设 $f(x), g(x)$ 的首项分别是 a_nx^n, b_mx^m.于是 $a_nb_m^{-1}x^{n-m}g(x)$ 的首项是 a_nx^n.令

$$f_1(x) = f(x) - a_nb_m^{-1}x^{n-m}g(x), \tag{4}$$

则 $\deg f_1(x) < n$.根据归纳假设,存在 $h_1(x), r_1(x)\in K[x]$,使得

$$f_1(x) = h_1(x)g(x) + r_1(x),\quad \deg r_1(x) < \deg g(x). \tag{5}$$

把(5)式代入(4)式,得

$$\begin{aligned} f(x) &= f_1(x) + a_nb_m^{-1}x^{n-m}g(x) \\ &= [h_1(x) + a_nb_m^{-1}x^{n-m}]g(x) + r_1(x). \end{aligned} \tag{6}$$

令　$h(x) = h_1(x) + a_nb_m^{-1}x^{n-m}$,则

$$f(x) = h(x)g(x) + r_1(x),\quad \deg r_1(x) < \deg g(x). \tag{7}$$

根据数学归纳法原理,定理 3 的存在性部分得证.

唯一性.设 $h(x), r(x), h'(x), r'(x)\in K[x]$,使得

$$f(x) = h(x)g(x) + r(x),\quad \deg r(x) < \deg g(x), \tag{8}$$

$$f(x) = h'(x)g(x) + r'(x),\quad \deg r'(x) < \deg g(x) \tag{9}$$

则从(8),(9)两式得

$$h(x)g(x) + r(x) = h'(x)g(x) + r'(x),$$

即

$$[h(x) - h'(x)]g(x) = r'(x) - r(x). \tag{10}$$

于是

$$\begin{aligned} \deg[h(x) - h'(x)] + \deg g(x) &= \deg[r'(x) - r(x)] \\ &\leq \max\{\deg r'(x), \deg r(x)\} < \deg g(x). \end{aligned} \tag{11}$$

假如　$h(x)\neq h'(x)$,则从(11)式得

$$\deg[h(x) - h'(x)] < 0,$$

矛盾.因此　$h(x) = h'(x)$.从而 $r(x) = r'(x)$.唯一性得证.　x

(3) 式中的 $h(x)$称为 $g(x)$除 $f(x)$的商式，$r(x)$称为 $g(x)$除 $f(x)$的余式.

推论 4　设 $f(x), g(x) \in K[x]$，且 $g(x) \neq 0$，则 $g(x) \mid f(x)$当且仅当 $g(x)$除 $f(x)$的余式为零.

证明　$g(x) \mid f(x)$　　　$f(x) = h(x)g(x), h(x) \in K[x]$.

　　　　　　　　　　　$g(x)$除 $f(x)$的余式为 0.　　　x

利用带余除法可以证明：对于 $K[x]$中的多项式 $f(x), g(x)$，如果在 $K[x]$中，$g(x)$不能整除 $f(x)$，那么把数域 K 扩大成数域 F 后，在 $F[x]$中，$g(x)$仍然不能整除 $f(x)$.即我们有下述命题：

命题 5　设 $f(x), g(x) \in K[x]$，数域 F　K，则

在 $K[x]$中，$g(x) \mid f(x)$　　　在 $F[x]$中，$g(x) \mid f(x)$.

证明　必要性.设在 $K[x]$中，$g(x) \mid f(x)$，则存在 $h(x) \in K[x]$，使得 $f(x) = h(x)g(x)$.由于 K　F，因此 $f(x), g(x), h(x) \in F[x]$.从而在 $F[x]$中，$g(x) \mid f(x)$.

充分性.设在 $F[x]$中，$g(x) \mid f(x)$.

当 $g(x) \neq 0$ 时，在 $K[x]$中作带余除法，有 $h(x), r(x) \in K[x]$，使得

$$f(x) = h(x)g(x) + r(x), \quad \deg r(x) < \deg g(x). \tag{12}$$

由于 $f(x), g(x), h(x), r(x) \in F[x]$，因此(12)式也可看成是在 $F[x]$中的带余除法.由于在 $F[x]$中，$g(x) \mid f(x)$，因此根据推论 4 得，$r(x) = 0$.从而在 $K[x]$中，$g(x) \mid f(x)$.

当 $g(x) = 0$ 时，从 $g(x) \mid f(x)$得，$f(x) = 0$.从而在 $K[x]$中，有 $g(x) \mid f(x)$.　　　x

在带余除法中，可以采用下面例 1 中的格式来作.

例 1　用 $g(x)$除 $f(x)$，求商式与余式，其中

$$f(x) = 2x^3 + 3x^2 + 5, \quad g(x) = x^2 + 2x - 1.$$

解

x^2+2x-1	$2x^3+3x^2\qquad+5$	$2x-1$
	$2x^3+4x^2-2x$	
	$-x^2+2x+5$	
	$-x^2-2x+1$	
	$4x+4$	

因此

$$2x^3+3x^2+5=(2x-1)(x^2+2x-1)+(4x+4),$$

即用 $g(x)$ 去除 $f(x)$，商式是 $2x-1$，余式是 $4x+4$.

在带余除法中，如果除式是一次多项式，则可以采用下述综合除法[①].

例2　设 $f(x)=2x^4-6x^3+3x^2-2x+5$，$g(x)=x-2$，求 $g(x)$ 除 $f(x)$ 所得的商式与余式.

解

$$\begin{array}{rrrrr|r} 2 & -6 & 3 & -2 & 5 & 2 \\ & 4 & -4 & -2 & -8 & \\ \hline 2 & -2 & -1 & -4 & -3 & \end{array}$$

因此　$f(x)=(2x^3-2x^2-x-4)(x-2)-3$. 即用 $x-2$ 去除 $f(x)$，商式是 $2x^3-2x^2-x-4$，余式是 -3.

习　题　7.2

1. 证明整除关系的传递性，即在 $K[x]$ 中，如果 $f(x)\mid g(x)$ 且 $g(x)\mid h(x)$，则 $f(x)\mid h(x)$.

2. 证明本节的命题2，即在 $K[x]$ 中，如果

$$g(x)\mid f_i(x),\quad i=1,2,\cdots,s,$$

则对于任意 $u_i(x)\in K[x]$，$i=1,2,\cdots,s$，有

$$g(x)\mid(u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots+u_s(x)f_s(x)).$$

3. 用 $g(x)$ 除 $f(x)$，求商式与余式.

(1) $f(x)=x^4-3x^2-2x-1$，$g(x)=x^2-2x+5$；

(2) $f(x)=x^4+x^3-2x+3$，$g(x)=3x^2-x+2$.

4. 设 $f(x)=x^4-3x^3+a_1x+a_0$，$g(x)=x^2-3x+1$，求 $g(x)$ 整除 $f(x)$ 的充分必要条件.

5. 用综合除法求一次多项式 $g(x)$ 除 $f(x)$ 所得的商式与余式.

(1) $f(x)=3x^4-5x^2+2x-1$，$g(x)=x-4$；

(2) $f(x)=5x^3-3x+4$，　$g(x)=x+2$.

*6. 设 $a,b\in Z$，如果有 $h\in Z$ 使得

$$a=hb,$$

---

① 关于综合除法的理论根据，可参看丘维声编著《高等代数(下册)》(高等教育出版社1996年12月第1版)第23页.

则称 b 整除 a,记作 $b|a$,此时 b 叫作 a 的因数(或因子),a 叫作 b 的倍数;否则,称 b 不能整除 a,记作 $b\nmid a$.证明:

(1) 如果 $a|b$ 且 $b|a$(此时称 a 与 b 相伴),则 $a=\pm b$;反之也成立;

(2) 如果 $a|b$ 且 $b|c$,则 $a|c$;

(3) 如果 $b|a_i$, $i=1,2,\cdots,s$,则对于任意 $u_i\in Z$, $i=1,2,\cdots,s$,有

$$b|(u_1a_1+u_2a_2+\cdots+u_sa_s).$$

(4) 如果 $b|a$,且 $a\neq 0$,则 $|b|\leq|a|$.

阅读材料一:整数环中的带余除法

定理 1　任给 $a,b\in Z$, $b\neq 0$,则存在唯一的一对整数 q,r,使得

$$a=bq+r,0\leq r<|b|.$$

证明　考虑集合

$$S=\{a-bs|s\in Z,a-bs\geq 0\}.$$

若 $b>0$,则 $a-b(-a^2)=a+a^2b\geq 0$;若 $b<0$,则 $a-ba^2\geq 0$.因此 S 非空,从而 S 至少包含一个最小元素,设为 $a-bq$.令 $r=a-bq$,根据 S 的定义,$r\geq 0$.如果 $r\geq|b|$,且 $b>0$,则 $r-b\geq 0$,从而 $a-bq-b\geq 0$,即 $a-b(q+1)\geq 0$.于是 $a-b(q+1)\in S$,即 $r-b\in S$,但 $r-b<r$,这与 r 的定义矛盾.如果 $r\geq|b|$,且 $b<0$,则 $0\leq r+b=a-b(q-1)$.于是 $r+b\in S$,但 $r+b<r$,矛盾.所以 $r<|b|$.存在性得证.

唯一性.假如还有整数 q',r',使得

$$a=bq'+r',0\leq r'<|b|,$$

则得 $bq+r=bq'+r'$.不妨设 $r\leq r'$.于是 $b(q-q')=r'-r$.因为 $0\leq r<|b|$, $0\leq r'<|b|$, $r\leq r'$,所以 $0\leq r'-r<|b|$.于是

$$|q-q'|=\frac{r'-r}{|b|}<1,$$

由此得 $q-q'=0$,即 $q=q'$.从而得 $r=r'$.　　x

§3　最大公因式

从上一节知道,数域 K 上的一元多项式环 $K[x]$ 有带余除法,这是 $K[x]$ 的一个重要性质.这一节我们要由此出发推导出 $K[x]$ 的另一个重要性质:$K[x]$ 中任何两个多项式都有最大公因式,并且 $f(x)$ 与 $g(x)$ 的最大公因式可以表成 $f(x)$ 与 $g(x)$ 的一个组合.

在 $K[x]$中,如果 $c(x)$既是 $f(x)$的因式,又是 $g(x)$的因式,则称 $c(x)$是 $f(x)$与 $g(x)$的一个公因式.在 $f(x)$与 $g(x)$的所有公因式中,具有下述性质的公因式特别重要,即

定义 1　$K[x]$中多项式 $f(x)$与 $g(x)$的一个公因式 $d(x)$如果具有下述性质;对于 $f(x)$与 $g(x)$的任一公因式 $c(x)$,都有 $c(x)\mid d(x)$,则称 $d(x)$是 $f(x)$与 $g(x)$的一个最大公因式.

从定义立即得出,两个零多项式的最大公因式就是 0;对于任意多项式 $f(x)$, $f(x)$是 $f(x)$与 0 的一个最大公因式.

对于 $K[x]$中任意两个多项式,是否存在它们的最大公因式?如果存在,如何找出它们的最大公因式?对于给定的两个多项式 $f(x)$与 $g(x)$,它们的最大公因式是否唯一?这些就是本节要讨论的问题.

我们先指出几个简单而有用的结论:

命题 1　设 $f(x), g(x), p(x), q(x)$都是 $K[x]$中的多项式,如果 $f(x)$与 $g(x)$的所有公因式组成的集合等于 $p(x)$与 $q(x)$的所有公因式组成的集合,那么 $f(x)$与 $g(x)$的最大公因式的集合等于 $p(x)$与 $q(x)$的最大公因式的集合.

证明　设 $d(x)$是 $f(x)$与 $g(x)$的一个最大公因式,则 $d(x)$是 $p(x)$与 $q(x)$的一个公因式.任取 $p(x)$与 $q(x)$的一个公因式 $\phi(x)$,则 $\phi(x)$也是 $f(x)$与 $g(x)$的公因式,从而 $\phi(x)\mid d(x)$.所以 $d(x)$是 $p(x)$与 $q(x)$的一个最大公因式.同理, $p(x)$与 $q(x)$的任一最大公因式也是 $f(x)$与 $g(x)$的最大公因式.　x

推论 2　设 $f(x), g(x)\in K[x]$, a, b是 K 中非零数,则 $f(x)$与 $g(x)$的最大公因式的集合等于 $af(x)$与 $bg(x)$的最大公因式的集合.

证明　显然 $f(x)$与 $g(x)$的任一公因式是 $af(x)$与 $bg(x)$的公因式.容易说明 $af(x)$与 $bg(x)$的任一公因式也是 $f(x)$与 $g(x)$的公因式.于是由命题 1 立即得出结论.　x

引理 1　在 $K[x]$中,如果有等式

$$f(x) = h(x)g(x) + r(x) \tag{1}$$

成立,则 $f(x)$与 $g(x)$的最大公因式的集合等于 $g(x)$与 $r(x)$的最大公因式的集合.

证明　设 $d(x)$是 $f(x)$与 $g(x)$的一个公因式,则 $d(x)\mid f(x)$,并且 $d(x)\mid g(x)$.因为从(1)式得

$$r(x) = f(x) - h(x)g(x),$$

所以 $d(x)\mid r(x)$.即 $d(x)$是 $g(x)$与 $r(x)$的一个公因式.现在任取 $g(x)$与 $r(x)$的一个公因式 $c(x)$,由(1)式得, $c(x)\mid f(x)$.于是 $c(x)$是 $f(x)$与 $g(x)$

的公因式.由命题1立即得到所要求的结论. x

现在来证明这一节的主要结果:

定理3 对于 $K[x]$中任意两个多项式 $f(x)$与 $g(x)$,存在它们的一个最大公因式 $d(x)$,并且 $d(x)$可以表示成 $f(x)$与 $g(x)$的一个组合,即有 $K[x]$中多项式 $u(x)$与 $v(x)$,使

$$d(x)=u(x)f(x)+v(x)g(x). \tag{2}$$

证明 如果 $g(x)=0$,则 $f(x)$就是 $f(x)$与 $g(x)$的一个最大公因式,并且

$$f(x)=1\cdot f(x)+1\cdot 0.$$

现在设 $g(x)\neq 0$.据带余除法,有 $h_1(x),r_1(x)\in K[x]$,使

$$f(x)=h_1(x)g(x)+r_1(x),\quad \deg r_1(x)<\deg g(x).$$

如果 $r_1(x)\neq 0$,则用 $r_1(x)$去除 $g(x)$,有 $h_2(x),r_2(x)\in K[x]$,使得

$$g(x)=h_2(x)r_1(x)+r_2(x),\quad \deg r_2(x)<\deg r_1(x).$$

又如果 $r_2(x)\neq 0$,则用 $r_2(x)$去除 $r_1(x)$,有 $h_3(x),r_3(x)\in K[x]$,使得

$$r_1(x)=h_3(x)r_2(x)+r_3(x),\quad \deg r_3(x)<\deg r_2(x).$$

如此辗转相除下去,显然,所得余式的次数不断降低,因此在有限次之后,必然有余式为零,即

$$r_2(x)=h_4(x)r_3(x)+r_4(x),\quad \deg r_4(x)<\deg r_3(x),$$

…………

$$r_{i-2}(x)=h_i(x)r_{i-1}(x)+r_i(x),\quad \deg r_i(x)<\deg r_{i-1}(x),$$

…………

$$r_{s-3}(x)=h_{s-1}(x)r_{s-2}(x)+r_{s-1}(x),\quad \deg r_{s-1}(x)<\deg r_{s-2}(x),$$

$$r_{s-2}(x)=h_s(x)r_{s-1}(x)+r_s(x),\quad \deg r_s(x)<\deg r_{s-1}(x),$$

$$r_{s-1}(x)=h_{s+1}(x)r_s(x)+0,$$

其中所有 $h_i(x),r_i(x)\in K[x]$.从上述等式的最后一个式子得出,$r_s(x)$是 $r_{s-1}(x)$与 $r_s(x)$的一个最大公因式.据引理1得,$r_s(x)$也是 $r_{s-2}(x)$与 $r_{s-1}(x)$的一个最大公因式;从而 $r_s(x)$也是 $r_{s-3}(x)$与 $r_{s-2}(x)$的一个最大公因式;依次往上推,$r_s(x)$也是 $f(x)$与 $g(x)$的一个最大公因式.这证明了:在对 $f(x)$与 $g(x)$作辗转相除时,最后一个不等于零的余式是 $f(x)$与 $g(x)$的一个最大公因式.

从上述等式中倒数第二个得

$$r_s(x)=r_{s-2}(x)-h_s(x)r_{s-1}(x),$$

再由倒数第三式得

$$r_{s-1}(x)=r_{s-3}(x)-h_{s-1}(x)r_{s-2}(x),$$

代入上式得

$$r_s(x) = (1 + h_s(x)h_{s-1}(x))r_{s-2}(x) - h_s(x)r_{s-3}(x).$$

同理用更上面的等式逐个地消去 $r_{s-2}(x), r_{s-3}(x), \cdots, r_1(x)$,可得

$$r_s(x) = u(x)f(x) + v(x)g(x),$$

其中 $u(x), v(x) \in K[x]$. x

定理 3 的证明给出了求两个多项式的最大公因式的方法,称它为辗转相除法.

任意给定 $K[x]$中两个多项式 $f(x)$与 $g(x)$,它们的最大公因式是否唯一?设 $d_1(x), d_2(x)$都是 $f(x)$与 $g(x)$的最大公因式,据定义得, $d_1(x) | d_2(x)$且 $d_2(x) | d_1(x)$.因此 $d_1(x)$与 $d_2(x)$相伴,即 $d_1(x)$与 $d_2(x)$仅相差一个非零数因子.这说明:两个多项式的最大公因式在相伴的意义下是唯一确定的.容易看出,两个不全为零的多项式的最大公因式一定是非零多项式,在这个情形,我们约定,用

$$(f(x), g(x))$$

来表示首项系数是 1 的那个最大公因式.

由推论 2 得出,设 $f(x), g(x)$不全为零,则对于 $a, b \in K$ 且 $a \neq 0, b \neq 0$,有

$$(f(x), g(x)) = (af(x), bg(x)).$$

例 1　设

$$f(x) = x^3 + x^2 - 7x + 2, \quad g(x) = 3x^2 - 5x - 2,$$

求$(f(x), g(x))$,并且把它表示成 $f(x)$与 $g(x)$的一个组合.

解　根据上面所述,在作辗转相除法时,可以用适当的非零数去乘被除式或者除式,以便使计算简单一些.

	$g(x)$	$3f(x)$	
$h_2(x) = 3x + 1$	$3x^2 - 5x - 2$	$3x^3 + 3x^2 - 21x + 6$	$x + \frac{8}{3} = h_1(x)$
	$3x^2 - 6x$	$3x^3 - 5x^2 - 2x$	
	$x - 2$	$8x^2 - 19x + 6$	
	$x - 2$	$8x^2 - \frac{40}{3}x - \frac{16}{3}$	
	0	$r_1(x) = -\frac{17}{3}x + \frac{34}{3}$	
		$-\frac{3}{17}r_1(x) = x - 2$	

因为最后一个不等于零的余式是 $r_1(x)$,所以

$$(f(x), g(x)) = x - 2.$$

把上述辗转相除过程写出来就是

$$3f(x) = \quad x + \frac{8}{3} \quad g(x) + r_1(x),$$

$$g(x) = (3x + 1) \quad -\frac{3}{17} r_1(x) \quad + 0.$$

于是

$$(f(x), g(x)) = -\frac{3}{17} r_1(x)$$

$$= -\frac{3}{17} \quad 3f(x) - \quad x + \frac{8}{3} \quad g(x)$$

$$= -\frac{9}{17} f(x) + \frac{1}{17}(3x + 8) g(x).$$

一个极端情形是两个多项式的最大公因式是非零数(即零次多项式),这个情形很重要,我们专门来讨论它.

定义 2　$K[x]$中的两个多项式 $f(x)$, $g(x)$,如果$(f(x), g(x)) = 1$,则称 $f(x)$与 $g(x)$互素.

显然,如果两个多项式互素,那么它们除去零次多项式外没有其他的公因式,反之亦然.

下面我们给出两个多项式互素的一个充分必要条件,它非常有用.

定理 4　$K[x]$中两个多项式 $f(x)$与 $g(x)$互素的充分必要条件是有 $K[x]$中的多项式 $u(x)$, $v(x)$,使

$$u(x)f(x) + v(x)g(x) = 1. \tag{3}$$

证明　必要性.从定理 3 立即得出.

现在证充分性.设(3)式成立.因为$(f(x), g(x)) \mid f(x)$,并且$(f(x), g(x)) \mid g(x)$,所以从(3)式得,$(f(x), g(x)) \mid 1$.于是

$$(f(x), g(x)) = 1.$$

x

利用定理 4 可以证明关于互素的多项式的一些重要性质:

性质 1　在 $K[x]$中,如果

$$f(x) \mid g(x)h(x), 且 (f(x), g(x)) = 1,$$

则

$$f(x) \mid h(x).$$

证明　若 $h(x) = 0$,则 $f(x) \mid h(x)$.下面设 $h(x) \neq 0$.因为$(f(x), g(x)) = 1$,所以有 $u(x), v(x) \in K[x]$,使得

$$u(x)f(x) + v(x)g(x) = 1$$

等式两边乘以 $h(x)$,得

$$u(x)f(x)h(x) + v(x)g(x)h(x) = h(x), \tag{4}$$

因为 $f(x) \mid g(x)h(x)$,所以 $f(x)$整除(4)式左端,从而

$$f(x) \mid h(x).$$

x

性质 2　在 $K[x]$中,如果

$$f(x)\mid h(x),\quad g(x)\mid h(x),\quad 且 (f(x),g(x))=1,$$

则
$$f(x)g(x)\mid h(x).$$

证明　因为 $f(x)\mid h(x)$,所以有 $p(x)\in K[x]$,使

$$h(x)=p(x)f(x). \tag{5}$$

因为 $g(x)\mid h(x)$,所以 $g(x)\mid p(x)f(x)$.因为 $(g(x),f(x))=1$,所以 $g(x)\mid p(x)$.于是有 $q(x)\in K[x]$,使得 $p(x)=q(x)g(x)$.代入(5)式得

$$h(x)=q(x)g(x)f(x),$$

所以 $f(x)g(x)\mid h(x)$.　x

性质3　在 $K[x]$中,如果

$$(f(x),h(x))=1,\quad (g(x),h(x))=1,$$

则
$$(f(x)g(x),h(x))=1$$

证明　因为$(f(x),h(x))=1$,$(g(x),h(x))=1$,所以有 $K[x]$中的多项式 $u_i(x)$, $v_i(x)$, $i=1,2$,使得

$$u_1(x)f(x)+v_1(x)h(x)=1,$$
$$u_2(x)g(x)+v_2(x)h(x)=1.$$

将上面两个等式相乘,得

$$u_1(x)u_2(x)f(x)g(x)+[u_1(x)f(x)v_2(x)$$
$$+v_1(x)u_2(x)g(x)+v_1(x)v_2(x)h(x)]h(x)=1,$$

据定理4得,$(f(x)g(x),h(x))=1$.　x

最大公因式和互素的概念可以推广到 $n(n>2)$个多项式的情形.

定义3　在 $K[x]$中,如果多项式 $c(x)$能整除多项式 $f_1(x)$, $f_2(x)$,…, $f_n(x)$的每一个,那么 $c(x)$叫做这 n 个多项式的一个公因式.设 $d(x)$是 $f_1(x)$,…, $f_n(x)$的一个公因式,如果 $d(x)$具有下述性质:$f_1(x)$,…, $f_n(x)$的每一个公因式都能整除 $d(x)$,则 $d(x)$称为 $f_1(x)$,…, $f_n(x)$的一个最大公因式.

用数学归纳法可以证明,在 $K[x]$中,任意 n个$(n\geq 2)$多项式 $f_1(x)$,…, $f_n(x)$的最大公因式存在,并且如果 $d_1(x)$是 $f_1(x)$,…, $f_{n-1}(x)$的一个最大公因式,则 $d_1(x)$与 $f_n(x)$的最大公因式是 $f_1(x)$,…, $f_{n-1}(x)$, $f_n(x)$的最大公因式.因此可以逐次用辗转相除法求出 n 个多项式的一个最大公因式.

从定义3立即得出,n 个多项式 $f_1(x)$,…, $f_n(x)$的最大公因式在相伴的意义下是唯一的.对于 n 个不全为零的多项式 $f_1(x)$,…, $f_n(x)$,我们约定,用

$$(f_1(x),f_2(x),\cdots,f_n(x))$$

表示首项系数是1的那个最大公因式.由上面一段知

$$(f_1(x),f_2(x),\cdots,f_n(x))=((f_1(x),\cdots,f_{n-1}(x)),f_n(x)). \tag{6}$$

从(6)式以及定理3得,有 $K[x]$中多项式 $u_i(x), i=1,\cdots,n$,使得

$$u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots+u_n(x)f_n(x)$$
$$=(f_1(x),f_2(x),\cdots,f_n(x)). \tag{7}$$

定义4　$K[x]$中 $n(n\geq 2)$个多项式 $f_1(x),f_2(x),\cdots,f_n(x)$,如果 $(f_1(x),f_2(x),\cdots,f_n(x))=1$,则称 $f_1(x),f_2(x),\cdots,f_n(x)$互素.

与定理4一样可以证明:在 $K[x]$中,n个多项式 $f_1(x),f_2(x),\cdots,f_n(x)$互素的充分必要条件是有 $K[x]$中多项式 $u_1(x),u_2(x),\cdots,u_n(x)$,使得

$$u_1(x)f_1(x)+u_2(x)f_2(x)+\cdots+u_n(x)f_n(x)=1. \tag{8}$$

注意: $n(n>2)$个多项式互素时,它们不一定两两互素.例如,设

$$f_1(x)=x+1,\ f_2(x)=x^2+3x+2,\ f_3(x)=x-1.$$

容易算出

$$(f_1(x),f_2(x))=x+1,(f_1(x),f_2(x),f_3(x))=1,$$

因此,$f_1(x),f_2(x),f_3(x)$互素,但是 $f_1(x)$与 $f_2(x)$不互素.

我们还要指出一点,设 K 与 F 都是数域,并且 K　F.设 $f(x)$, $g(x)\in K[x]$,也可以把 $f(x)$与 $g(x)$看成 $F[x]$中的多项式.注意 $f(x)$与 $g(x)$在 $K[x]$中的公因式与它们在 $F[x]$中的公因式不一定相同,例如设

$$f(x)=x^2+1,\quad g(x)=x^3+x^2+x+1,$$

则 $f(x)$与 $g(x)$在 $R[x]$中没有一次公因式,但是 $f(x)$与 $g(x)$在 $C[x]$中有一次公因式 $x+i$ 与 $x-i$.容易看出它们在 $R[x]$中的最大公因式是 x^2+1,在 $C[x]$中的最大公因式也是 x^2+1.一般地,我们有下述结论:

命题5　设数域 F　K,则对于 $K[x]$中任意两个多项式 $f(x)$与 $g(x)$,它们在 $K[x]$中的首项系数为1的最大公因式与它们在 $F[x]$中的首项系数为1的最大公因式相同,也就是说,当数域扩大时,$f(x)$与 $g(x)$的首项系数为1的最大公因式不改变.

证明　若 $f(x)=g(x)=0$,则 $f(x)$与 $g(x)$在 $K[x]$中的最大公因式是零多项式,在 $F[x]$中的最大公因式也是零多项式.

下面设 $f(x)$与 $g(x)$不全为零.设 $d_1(x)$是 $f(x)$与 $g(x)$在 $K[x]$中的首项系数为1的最大公因式,设 $d_2(x)$是 $f(x)$与 $g(x)$在 $F[x]$中的首项系数为1的最大公因式.在 $K[x]$中对 $f(x)$与 $g(x)$做辗转相除法,设最后一个不等于零的余式是 $r_s(x)$,其首项系数为 c,则 $d_1(x)=\dfrac{1}{c}r_s(x)$,由于每一步带余除法也可看成是在 $F[x]$中进行的(根据带余除法的唯一性),因此 $r_s(x)$也是 $f(x)$与 $g(x)$在 $F[x]$中的一个最大公因式,从而

$$d_2(x)=\frac{1}{c}r_s(x)=d_1(x).$$

x

推论 6　设数域 F　K.设 $f(x),g(x)\in K[x]$,则 $f(x)$与 $g(x)$在 $K[x]$中互素,当且仅当 $f(x)$与 $g(x)$在 $F[x]$中互素.即互素性不随数域的扩大而改变.

证明　　$f(x)$与 $g(x)$在 $K[x]$中互素

在 $K[x]$中,$(f(x),g(x))=1$

在 $F[x]$中,$(f(x),g(x))=1$

$f(x)$与 $g(x)$在 $F[x]$中互素.　　x

习　题　7.3

1. 求 $f(x)$与 $g(x)$的最大公因式,并且把$(f(x),g(x))$表示成 $f(x)$与 $g(x)$的一个组合:

(1) $f(x)=x^4+3x^3-x^2-4x-3$,

$g(x)=3x^3+10x^2+2x-3$;

(2) $f(x)=x^4+6x^3-6x^2+6x-7$,

$g(x)=x^3+x^2-7x+5$.

2. 证明:在 $K[x]$中,如果 $d(x)$是 $f(x)$与 $g(x)$的一个组合,并且 $d(x)$是 $f(x)$与 $g(x)$的一个公因式,则 $d(x)$是 $f(x)$与 $g(x)$的一个最大公因式.

3. 证明:在 $K[x]$中,$(f,g)h$ 是 fh 与 gh 的一个最大公因式;特别地,若 $h(x)$的首项系数为 1,则

$$(fh,gh)=(f,g)h.$$

4. 证明:在 $K[x]$中,如果 $f(x),g(x)$不全为零,则

$$\frac{f(x)}{(f(x),g(x))},\frac{g(x)}{(f(x),g(x))}=1.$$

5. 证明:在 $K[x]$中,如果 $f(x),g(x)$不全为零,并且

$$u(x)f(x)+v(x)g(x)=(f(x),g(x)),$$

则$(u(x),v(x))=1$.

6. 证明:在 $K[x]$中,如果$(f,g)=1$,那么

$$(fg,f+g)=1.$$

7. 设 $f(x),g(x)\in K[x]$,并且 $a,b,c,d\in K$,使得

$$ad-bc\neq 0.$$

证明:$(af+bg,cf+dg)=(f,g)$.

8. 证明:在 $K[x]$中,如果$(f(x),g(x))=1$,则对任意正整数 m,有

$$(f(x^m),g(x^m))=1.$$

9. 证明:$K[x]$中两个非零多项式 $f(x)$与 $g(x)$不互素的充分必要条件是,存在两个非零多项式 $u(x),v(x)$,使

$$u(x)f(x)=v(x)g(x),$$
$$\deg u(x)<\deg g(x),\quad \deg v(x)<\deg f(x).$$

10. 设 $f(x),g(x)\in K[x]$, $K[x]$中的一个多项式 $m(x)$称为 $f(x)$与 $g(x)$的一个最小公倍式,如果

1) $f(x)\mid m(x), g(x)\mid m(x)$;

2) $f(x)$与 $g(x)$的任一公倍式(即 $K[x]$中既能被 $f(x)$整除,又能被 $g(x)$整除的多项式)都是 $m(x)$的倍式.

(1) 证明 $K[x]$任意两个多项式都有最小公倍式,并且在相伴的意义下是唯一的;

(2) 用$[f(x),g(x)]$表示首项系数是 1 的那个最小公倍式.证明:如果 $f(x),g(x)$的首项系数都是 1,则

$$[f(x),g(x)]=\frac{f(x)g(x)}{(f(x),g(x))}.$$

*11. 设 $A\in M_n(K)$, $f(x),g(x)\in K[x]$.证明:如果 $d(x)$是 $f(x)$与 $g(x)$的一个最大公因式,则齐次线性方程组 $d(A)X=0$的解空间等于 $f(A)X=0$ 的解空间与 $g(A)X=0$ 的解空间的交.

*12. 设 $A\in M_n(K)$, $f_1(x), f_2(x)\in K[x]$,记 $f(x)=f_1(x)f_2(x)$.证明:如果 $(f_1(x),f_2(x))=1$,则 $f(A)X=0$ 的任一个解可以唯一地表示成 $f_1(A)X=0$ 的一个解与 $f_2(A)X=0$的一个解的和.

阅读材料二:最大公因数

在整数环 Z 中也有带余除法,从而可以证明任意两个整数都有最大公因数.先给出定义.

定义 1 设 a,b是整数,若整数 c既是 a的因数,又是 b的因数,则称 c是 a与 b的一个公因数(或公因子).设 d是 a与 b的一个公因数,如果 a与 b的每一个公因数都能整除 d,则称 d是 a与 b的一个最大公因数(或最大公因子).

引理 1 在 Z 中,如果有等式

$$a=bq+c$$

成立,则 a与 b的最大公因数是 b与 c的最大公因数,反之亦然.

证明 类似于§3 的引理 1 的证明. x

定理 1 任意给定两个整数 a,b,都存在它们的一个最大公因数 d,并且 d可以表示成 a与 b的一个组合,即存在整数 u,v,使得

$$d=ua+vb. \tag{1}$$

证明 如果 $b=0$,则 a就是 a与 b的一个最大公因数,并且

$$a=1\cdot a+1\cdot 0.$$

现在设 $b\neq 0$.容易看出 a与 b的最大公因数是 a与 $-b$的最大公因数,反

之亦然.因此我们不妨设 $b>0$.据带余除法,有整数 q_1, r_1,使得

$$a=bq_1+r_1, 0\leq r_1<b.$$

如果 $r_1\neq 0$,则用 r_1 去除 b,有整数 q_2, r_2,使得

$$b=r_1q_2+r_2, 0\leq r_2<r_1.$$

如果 $r_2\neq 0$,再用 r_2 去除 r_1.如此辗转相除下去,显然所得余数(正整数)不断减小,因此在有限次后,必然有余数为零.以下类似于定理3的相应部分的证明,可证出最后一个不等于零的余数就是 a 与 b 的一个最大公因数,并且它可以表示成 a 与 b 的一个组合. x

从定义1可知,a 与 b 的最大公因数在相伴的意义下(即相差一个正负号)是唯一的.对于不全为零的整数 a, b,我们约定用 (a, b) 表示正的那个最大公因数.

定义2　设 $a, b\in \mathbf{Z}$,如果 $(a, b)=1$,则称 a 与 b 互素.

显然,如果两个整数互素,那么它们除去±1以外没有其他公因数,反之亦然.

定理2　两个整数 a 与 b 互素的充分必要条件是有整数 u, v,使得

$$ua+vb=1.$$

证明　必要性从定理1立即得出.充分性易证. x

从定理2可以得出

定理3　在 $\mathbf{Z}$ 中,如果 $a\mid bc$,并且 $(a, b)=1$,则 $a\mid c$.

证明留给读者.

定义3　在 $\mathbf{Z}$ 中,如果 c 是 $a_1, a_2, \cdots, a_n$ 中每一个的因数,则称 c 是 $a_1, a_2, \cdots, a_n$ 的一个公因数.设 d 是 $a_1, a_2, \cdots, a_n$ 的一公因数,如果 $a_1, a_2, \cdots, a_n$ 的每一个公因数都能整除 d,则称 d 是 $a_1, a_2, \cdots, a_n$ 的一个最大公因数.

用数学归纳法可以证明,任意 $n(n\geq 2)$ 个整数都有最大公因数,并且如果 d_1 是 $a_1, a_2, \cdots, a_{n-1}$ 的一个最大公因数,则 d_1 与 a_n 的最大公因数是 $a_1, a_2, \cdots, a_{n-1}, a_n$ 的最大公因数.因此可以用辗转相除法求出 n 个整数的一个最大公因数.

从定义3得出,n 个整数 $a_1, a_2, \cdots, a_n$ 的最大公因数在相伴的意义下是唯一的,对于 n 个不全为零的整数 $a_1, a_2, \cdots, a_n$,我们约定用 $(a_1, a_2, \cdots, a_n)$ 表示正的那个最大公因数.从上面一段知

$$(a_1, a_2, \cdots, a_n)=((a_1, a_2, \cdots, a_{n-1}), a_n). \tag{2}$$

从(2)式以及定理1可得出,对于 n 个不全为零的整数 $a_1, a_2, \cdots, a_n$,存在 n 个整数 $u_1, u_2, \cdots, u_n$,使得

$$u_1a_1+u_2a_2+\cdots+u_na_n=(a_1, a_2, \cdots, a_n). \tag{3}$$

定义 4　对于 n 个整数 $a_1, a_2, \cdots, a_n$,如果 $(a_1, a_2, \cdots, a_n) = 1$,则称 $a_1, a_2, \cdots, a_n$ 互素.

n 个整数 $a_1, a_2, \cdots, a_n$ 互素的充分必要条件是有整数 $u_1, u_2, \cdots, u_n$,使得

$$u_1 a_1 + u_2 a_2 + \cdots + u_n a_n = 1.$$

定义 5　设 a, b 是整数,一个整数 m 称为 a 与 b 的最小公倍数,如果

1) $a \mid m, b \mid m$;

2) 从 $a \mid l, b \mid l$,可推出 $m \mid l$.

可以证明,任意两个整数都有最小公倍数,并且在相伴的意义下是唯一的.对于全不为零的整数 a, b,我们用 $[a, b]$ 表示正的那个最小公倍数.对于 $a > 0$, $b > 0$,有

$$[a, b](a, b) = ab. \tag{4}$$

证明留给读者.

类似于定义 5,可给出 $n(n > 2)$ 个整数的最小公倍数的定义,可以证明它们存在,并且在相伴意义下唯一.对于全不为零的整数 $a_1, a_2, \cdots, a_n$,用 $[a_1, a_2, \cdots, a_n]$ 表示正的那个最小公倍数,可证

$$[a_1, a_2, \cdots, a_n] = [[a_1, a_2, \cdots, a_{n-1}], a_n].$$

§4　不可约多项式,唯一因式分解定理

我们已经知道,数域 K 上的一元多项式环 $K[x]$ 中有带余除法.由此推导出,$K[x]$ 中任意两个多项式都有最大公因式.现在我们利用这些结论来研究 $K[x]$ 的结构.与整数环 Z 类比:每一个正整数都能表示成有限多个素数的乘积,我们自然可以问:$K[x]$ 中每一个多项式是否能表示成有限多个具有像"素数"那样性质的多项式的乘积?一个大于 1 的正整数 p 称为素数,如果 p 的正因子只有 1 和 p 自身.由此受到启发,引出下述概念:

定义 1　$K[x]$ 中一个次数大于零的多项式 $p(x)$,如果它在 $K[x]$ 中的因式只有零次多项式和 $p(x)$ 的相伴元,则称 $p(x)$ 是数域 K 上的一个不可约多项式;否则称 $p(x)$ 是可约多项式.

性质 1　$K[x]$ 中不可约多项式 $p(x)$ 与任一多项式 $f(x)$ 的关系只有两种可能:或者 $p(x) \mid f(x)$,或者 $p(x)$ 与 $f(x)$ 互素.

证明　由于 $(p(x), f(x))$ 是 $p(x)$ 的因式,而 $p(x)$ 不可约,因此 $(p(x), f(x))$ 是零次多项式,或者 $(p(x), f(x)) \sim p(x)$.从而 $(p(x), f(x)) = 1$,或者 $p(x) \mid (p(x), f(x))$.从后者据整除的传递性得出 $p(x) \mid f(x)$.　x

性质 2　在 $K[x]$ 中,如果 $p(x)$ 不可约,且 $p(x) \mid f(x)g(x)$,则 $p(x) \mid f(x)$ 或 $p(x) \mid g(x)$.

证明　如果 $p(x) \mid f(x)$,则结论成立.下面设 $p(x) \nmid f(x)$.由于 $p(x)$ 不可约,因此据性质 1 得,$(p(x), f(x)) = 1$.于是从 $p(x) \mid f(x)g(x)$ 得出,$p(x) \mid g(x)$. x

利用数学归纳法,性质 2 可以推广为:在 $K[x]$ 中,如果 $p(x)$ 不可约,且 $p(x) \mid f_1(x)f_2(x)\cdots f_s(x)$,则 $p(x) \mid f_j(x)$,对于某个 $j \in \{1, 2, \cdots, s\}$.

性质 3　$K[x]$ 中,$p(x)$ 不可约当且仅当 $p(x)$ 不能分解成两个次数较 $p(x)$ 的次数低的多项式的乘积.

证明　必要性.如果 $p(x)$ 不可约,则 $p(x)$ 的因式只有零次多项式和 $p(x)$ 的相伴元,因此 $p(x)$ 不能分解成两个次数较 $p(x)$ 的次数低的多项式的乘积.

充分性.用反证法,假如 $p(x)$ 可约,则 $p(x)$ 有因式 $g(x)$ 使得 $0 < \deg g(x) < \deg p(x)$.从而存在 $h(x) \in K[x]$ 使得

$$p(x) = h(x)g(x),$$

于是

$$\deg p(x) = \deg h(x) + \deg g(x).$$

由此推出,

$$0 < \deg h(x) < \deg p(x).$$

这与已知条件矛盾. x

从性质 3 立即得出,$K[x]$ 中的每一个 1 次多项式一定是不可约多项式.

从性质 3 可以猜测有下述结论:

定理 1(唯一因式分解定理)　$K[x]$ 中每一个次数大于零的多项式 $f(x)$ 都能唯一地分解成数域 K 上有限多个不可约多项式的乘积.所谓唯一性是指,如果 $f(x)$ 有两个这样的分解式:

$$f(x) = p_1(x)p_2(x)\cdots p_s(x) = q_1(x)q_2(x)\cdots q_t(x),$$

则一定有 $s = t$,且适当排列因式的次序后有

$$p_i(x) \sim q_i(x), \quad i = 1, 2, \cdots, s.$$

证明　先证分解式的存在性.对多项式的次数 n 作数学归纳法.

因为一次多项式都是不可约的,所以 $n = 1$ 时结论成立.

假设对于次数小于 n 的多项式结论成立,现在来看 n 次多项式 $f(x)$.

如果 $f(x)$ 是不可约多项式,则结论显然成立.如果 $f(x)$ 是可约多项式,则有

$$f(x) = f_1(x)f_2(x),$$

其中 $f_i(x) \in K[x]$,并且 $\deg f_i(x) < \deg f(x)$, $i = 1, 2$.由归纳假设,$f_1(x)$ 与 $f_2(x)$ 都可以分解成数域 K 上有限多个不可约多项式的乘积,把 $f_1(x)$ 与 $f_2(x)$ 的分解式合起来就得到 $f(x)$ 的一个分解式.

由归纳法原理,结论普遍成立.

现在证唯一性.假设 $f(x)$有两个这样的分解式

$$f(x)=p_1(x)p_2(x)\cdots p_s(x)=q_1(x)q_2(x)\cdots q_t(x), \tag{1}$$

其中 $p_i(x),q_j(x),i=1,\cdots,s,j=1,\cdots,t$ 都是数域 K 上的不可约多项式.

我们对分解式中不可约因式个数 s 作归纳法.

当 $s=1$ 时,$f(x)=p_1(x)$,则 $f(x)$是不可约多项式.由于不可约多项式的因式只有它的相伴元和零次多项式,所以从 $q_1(x)\mid f(x)$得

$$q_1(x)\sim f(x),$$

从而 $f(x)=cq_1(x),c\in K^*$.因此 $t=1$,且 $p_1(x)\sim q_1(x)$.

假设分解式中不可约因式的个数为 $s-1$ 时唯一性成立.现在来看不可约因式的个数为 s 的情形.

由(1)式,$p_1(x)\mid q_1(x)q_2(x)\cdots q_t(x)$,因此 $p_1(x)$必能整除其中的一个,不妨设

$$p_1(x)\mid q_1(x),$$

因为 $q_1(x)$也是不可约多项式,所以 $p_1(x)\sim q_1(x)$,即

$$p_1(x)=c_1q_1(x),\quad c_1\in K,\quad c_1\neq 0, \tag{2}$$

把(2)代入(1)式,并且两边消去 $q_1(x)$,得

$$p_2(x)\cdots p_s(x)=(c_1^{-1}q_2(x))\cdots q_t(x). \tag{3}$$

由归纳假设,有

$$s-1=t-1,\text{即 } s=t, \tag{4}$$

并且适当排列因式的次序之后有

$$p_2(x)\sim c_1^{-1}q_2(x),p_3(x)\sim q_3(x),\cdots,p_s(x)\sim q_s(x). \tag{5}$$

从而有

$$p_i(x)\sim q_i(x),\quad i=1,2,\cdots,s. \tag{6}$$

根据数学归纳法原理,唯一性得证. x

从唯一性的证明中可看出,$f(x)$的任一不可约因式一定与 $f(x)$的分解式中的某一个不可约因式相伴.因此 $f(x)$的分解式给出了它的全部不可约因式(在相伴意义下).

$K[x]$中的唯一因式分解定理在理论上非常重要,但是至今仍没有一个统一的方法来做因式分解,也就是没有统一的方法求出一个次数大于零的多项式的所有不可约因式.

在多项式 $f(x)$的分解式中,可以把每一个不可约因式的首项系数提出来,使它们成为首项系数为 1 的多项式,再把相同的不可约因式的乘积写成乘幂的形式,于是 $f(x)$的分解式成为

$$f(x) = cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_m^{r_m}(x), \tag{7}$$

其中 c 是 $f(x)$ 的首项系数, $p_1(x), p_2(x), \cdots, p_m(x)$ 是不同的首项系数为 1 的不可约多项式, $r_1, r_2, \cdots, r_m$ 是正整数. 这种分解式(7)称为 $f(x)$ 的标准分解式.

从理论研究的角度,如果已知两个多项式 $f(x)$ 与 $g(x)$ 的标准分解式,那么 $(f(x), g(x))$ 等于那些同时在 $f(x)$ 与 $g(x)$ 的标准分解式中出现的不可约多项式的方幂的乘积,所带的方幂的指数等于它在 $f(x)$ 与 $g(x)$ 中所带的方幂的指数中较小的一个. 即设

$$f(x) = ap_1^{k_1}(x)\cdots p_l^{k_l}(x)\,p_{l+1}^{k_{l+1}}(x)\cdots p_m^{k_m}(x), \tag{8}$$

$$g(x) = bp_1^{t_1}(x)\cdots p_l^{t_l}(x)\,q_{l+1}^{t_{l+1}}(x)\cdots q_s^{t_s}(x), \tag{9}$$

则

$$(f(x), g(x)) = p_1^{\min\{k_1, t_1\}}(x)\cdots p_l^{\min\{k_l, t_l\}}(x). \tag{10}$$

理由如下: 显然(10)式右端的多项式的乘积 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的一个公因式. 任取 $f(x)$ 与 $g(x)$ 的一个次数大于零的公因式 $c(x)$, 在 $c(x)$ 的标准分解式中任取一个不可约因式 $u(x)$, 由于 $u(x) \mid f(x)$ 且 $u(x) \mid g(x)$, 所以 $u(x)$ 等于某个 $p_i(x)$, 其中 $1 \leq i \leq l$; 并且 $u(x)$ 在 $c(x)$ 的标准分解式中的幂指数不超过 $\min\{k_i, t_i\}$. 于是得出, $c(x) \mid d(x)$. 这证明了 $d(x)$ 是 $f(x)$ 与 $g(x)$ 的最大公因式. x

由于把多项式分解成不可约因式的乘积没有统一的方法, 因此上述求最大公因式的方法不能代替辗转相除法.

我们举一个简单例子来看一下唯一因式分解定理的应用.

例 1　证明 $x^2 - 2$ 在有理数域上不可约.

证明　若 $x^2 - 2$ 在有理数域 Q 上可约, 则它的标准分解式为

$$x^2 - 2 = (x + a)(x + b), \quad a, b \in Q. \tag{11}$$

(11)式也可以看成 $x^2 - 2$ 在实数域 R 上的一个不可约因式分解. 另一方面我们知道 $x^2 - 2$ 在 R 上有如下的不可约因式分解

$$x^2 - 2 = (x + \ 2)(x - \ 2). \tag{12}$$

由 $R[x]$ 中唯一因式分解定理得

$$x + a = x + \ 2 \quad \text{或} \quad x + a = x - \ 2.$$

由此推出 $a = \ 2$ 或 $a = - \ 2$, 这与 $a \in Q$ 矛盾. 因此 $x^2 + 2$ 在 Q 上不可约.

请读者给出 $x^2 - 2$ 在有理数域上不可约的另一种证法. (提示: 仍用反证法, 从(11)式根据多项式相等的定义去推出矛盾.)

习 题 7.4

1. 证明多项式 x^2+1 在有理数域上不可约.

2. 分别在复数域,实数域和有理数域上分解下列多项式为不可约因式的乘积.

(1) x^4+1; (2) x^4+4.

3. 证明:在 $K[x]$ 中, $g^2(x)\mid f^2(x)$ 当且仅当 $g(x)\mid f(x)$.

4. 证明本节性质 2 的逆命题为真,即设 $p(x)$ 是 $K[x]$ 中一个次数大于零的多项式,如果对于任意 $f(x), g(x)\in K[x]$,从 $p(x)\mid f(x)g(x)$ 可以推出 $p(x)\mid f(x)$ 或者 $p(x)\mid g(x)$,那么 $p(x)$ 是不可约多项式.

*5. 证明:数域 K 上一个次数大于零的多项式 $f(x)$ 是 $K[x]$ 中某一不可约多项式的方幂的充分必要条件是对于任意 $g(x)\in K[x]$,必有 $(f(x), g(x))=1$,或者存在一个正整数 m,使得 $f(x)\mid g^m(x)$.

*6. 证明:数域 K 上一个次数大于零的多项式 $f(x)$ 是 $K[x]$ 中某一不可约多项式的方幂的充分必要条件是对于任意 $g(x), h(x)\in K[x]$,从 $f(x)\mid g(x)h(x)$ 可以推出 $f(x)\mid g(x)$,或者存在一个正整数 m,使得 $f(x)\mid h^m(x)$.

7. 在 $K[x]$ 中,设 $(f, g_i)=1$, $i=1,2$,证明:

$$(fg_1, g_2)=(g_1, g_2).$$

阅读材料三:整数环的唯一因子分解定理

定义 1 一个大于 1 的整数,如果它的正因数只有 1 和它本身,就叫做素数,否则就叫做合数.

从定义 1 立即得出,一个素数 p 与任一整数 a 之间的关系只有两种可能:或者 $(p, a)=1$,或者 $p\mid a$.

由上述关系,立即得出

定理 1 如果 p 是素数,那么对于任意整数 a, b,从 $p\mid ab$ 可以推出 $p\mid a$ 或者 $p\mid b$.

定理 1 可以推广为:如果素数 p 能整除一些整数 $a_1, a_2, \cdots, a_s$ 的乘积 $a_1a_2\cdots a_s$,则 p 一定整除其中的一个.

定理 1 的逆命题也成立.证明留给读者.

定理 2(整数的唯一因子分解定理) 任一大于 1 的整数 a 能唯一地分解成有限个素数的乘积.所谓唯一性是说,如果 a 有两个这样的分解式

$$a=p_1p_2\cdots p_s=q_1q_2\cdots q_t,$$

则一定有 $s=t$,并且适当排列因子的次序后有

$$p_i = q_i, \quad i=1,\cdots,s.$$

证明　先证分解式的存在性,用数学归纳法.

当 $a=2$ 时,2 是素数,因此结论显然成立.

假设对小于 a 且大于 1 的一切整数,结论成立,考虑 $a(a>1)$.如果 a 是素数,则结论成立.如果 a 是合数,则有两个大于 1 且小于 a 的整数 b,c,使得 $a=bc$.

由归纳假设,b 与 c 分别能分解成有限个素数乘积,从而 a 能分解成有限个素数的乘积.

由归纳法原理,存在性得证.

唯一性证明与§4 的定理 1 的唯一性证明类似,请读者完成.　　x

定理 2 称为算术基本定理,因为它在整数的整除性理论中起着基本重要的作用.算术基本定理告诉我们,任一大于 1 的整数 a 能唯一地写成(除了因子的排列次序外)

$$a = p_1^{r_1} p_2^{r_2} \cdots p_m^{r_m}, \tag{1}$$

其中 $p_1, p_2, \cdots, p_m$ 是不同的素数,$r_1, \cdots, r_m$ 是正整数,(1)式称为 a 的标准分解式.

如果已知两个大于 1 的整数 a, b 的标准分解式

$$a = p_1^{k_1} \cdots p_s^{k_s} p_{s+1}^{k_{s+1}} \cdots p_m^{k_m}, \tag{2}$$

$$b = p_1^{l_1} \cdots p_s^{l_s} q_{s+1}^{l_{s+1}} \cdots q_n^{l_n}, \tag{3}$$

则

$$(a, b) = p_1^{r_1} \cdots p_s^{r_s}, \ r_i = \min\{k_i, l_i\}, \ i=1,\cdots,s, \tag{4}$$

$$[a, b] = p_1^{\delta_1} \cdots p_s^{\delta_s} p_{s+1}^{k_{s+1}} \cdots p_m^{k_m} q_{s+1}^{l_{s+1}} \cdots q_n^{l_n}, \ \delta_i = \max\{k_i, l_i\}, \ i=1,\cdots,s. \tag{5}$$

(4)式与(5)式的证明留给读者作为练习.

§5　重　因　式

上一节我们已证明 $K[x]$中每一个次数大于零的多项式 $f(x)$能唯一地分解成数域 K 上有限多个不可约多项式的乘积.如果 $f(x)$的分解式中每一个不可约因式只出现 1 次,这种情形是特别重要的情形.这一节我们要给出识别这种情形的一个统一的方法.

定义 1　在 $K[x]$中,不可约多项式 $p(x)$称为多项式 $f(x)$的 k 重因式,如果 $p^k(x) \mid f(x)$,而 $p^{k+1}(x) \nmid f(x)$.

如果 $k=0$,则 $p(x) \nmid f(x)$,因此 $p(x)$不是 $f(x)$的因式;如果 $k=1$,则

$p(x)$称为 $f(x)$的单因式;如果 $k>1$,则$p(x)$称为 $f(x)$的重因式.

显然,如果 $f(x)$的标准分解式为

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_m^{r_m}(x),$$

则 $p_1(x),p_2(x),\cdots,p_m(x)$分别是 $f(x)$的 r_1 重,r_2 重,$\cdots$,r_m 重因式.指数 $r_i=1$ 的那些不可约因式是单因式;指数 $r_i>1$ 的那些不可约因式是重因式.所以 $f(x)$的分解式中每一个不可约因式只出现1次的情形也就是 $f(x)$没有重因式的情形.如何判别一个多项式有没有重因式呢?由于没有一般的方法来求一个多项式的标准分解式,因此我们必须寻找别的方法来判断一个多项式有没有重因式.

我们先看一个简单例子,以便从中受到启发.

设 $f(x)=(x+1)^3\in R[x]$,这时 $f(x)$有重因式.如果我们把 $f(x)$看成数学分析中讨论的多项式函数,那么对 $f(x)$可以求导数,得 $f'(x)=3(x+1)^2$.于是$(f(x),f'(x))=(x+1)^2$.从这个例子受到启发,有可能运用导数概念以及最大公因式的求法来讨论一个多项式有没有重因式的问题.由于我们现在讲的多项式是任意数域 K 上一个不定元的多项式,而数学分析中的多项式函数是实变量 x 的函数,其导数概念涉及极限概念,因此我们不能直接引用数学分析中多项式函数的导数概念,我们必须给任意数域 K 上一元多项式的导数下个定义,当然这个定义是从分析中多项式函数的导数公式受到启发的.

定义 2　对于 $K[x]$中的多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,$$

我们把 $K[x]$中的多项式

$$na_nx^{n-1}+(n-1)a_{n-1}x^{n-2}+\cdots+a_1,$$

叫做 $f(x)$的导数(或一阶导数),记作 $f'(x)$.

$f'(x)$的导数叫做 $f(x)$的二阶导数,记作 $f''(x)$;$f''(x)$的导数叫做 $f(x)$的三阶导数,记作 $f\ (x)$,等等.$f(x)$的 k阶导数也记作 $f^{(k)}(x)$.

从定义 2 立即得出,一个 n次多项式的导数是一个 $n-1$ 次多项式;它的 n 阶导数是 K 中一个非零数;它的 $n+1$ 阶导数等于零.零多项式的导数是零多项式.

根据定义 2,通过直接验证,可以得出 $K[x]$中多项式的导数的基本公式:

$$[f(x)+g(x)]'=f'(x)+g'(x),$$

$$[cf(x)]'=cf'(x),\ c\in K,$$

$$[f(x)g(x)]'=f'(x)g(x)+f(x)g'(x),$$

$$[f^m(x)]'=mf^{m-1}(x)f'(x).$$

仍看从前的简单例子.不可约多项式 $x+1$ 是 $f(x)=(x+1)^3$ 的 3 重因

式.由于按定义2和上述公式可得出，$f'(x)=3(x+1)^2$，因此 $x+1$ 是 $f'(x)$ 的2重因式.这个例子的结论有一般性，即

定理1　在 $K[x]$ 中，如果不可约多项式 $p(x)$ 是 $f(x)$ 的一个 $k(k\geq 1)$ 重因式，则 $p(x)$ 是 $f(x)$ 的导数 $f'(x)$ 的一个 $k-1$ 重因式.特别地，多项式 $f(x)$ 的单因式不是 $f(x)$ 的导数 $f'(x)$ 的因式.

证明　因为 $p(x)$ 是 $f(x)$ 的 k 重因式，所以存在 $g(x)\in K[x]$，使

$$f(x)=p^k(x)g(x),\quad p(x)\nmid g(x).$$

求 $f(x)$ 的导数，得

$$f'(x)=p^{k-1}(x)[kp'(x)g(x)+p(x)g'(x)].$$

因为 $p(x)\nmid kp'(x)$，$p(x)\nmid g(x)$，并且 $p(x)$ 是不可约多项式，所以 $p(x)\nmid kp'(x)g(x)$.但是 $p(x)\mid p(x)g'(x)$，所以 $p(x)$ 不能整除上述等式右端括号里的和.因此 $p(x)$ 是 $f'(x)$ 的 $k-1$ 重因式.　□

推论2　在 $K[x]$ 中，不可约多项式 $p(x)$ 是 $f(x)$ 的重因式的充分必要条件为 $p(x)$ 是 $f(x)$ 与 $f'(x)$ 的公因式.

证明　必要性.设不可约多项式 $p(x)$ 是 $f(x)$ 的 k 重因式，其中 $k>1$，则据定理1得，$p(x)$ 是 $f'(x)$ 的 $k-1$ 重因式，从而 $p(x)$ 是 $f(x)$ 与 $f'(x)$ 的公因式.

充分性.设不可约多项式 $p(x)$ 是 $f(x)$ 与 $f'(x)$ 的公因式.据定理1知，$p(x)$ 不是 $f(x)$ 的单因式，所以 $p(x)$ 是 $f(x)$ 的重因式.　□

从推论2立即得到：$K[x]$ 中次数大于零的多项式 $f(x)$ 有重因式的充分必要条件是，$f(x)$ 与 $f'(x)$ 有次数大于零的公因式.从而得出下述定理3：

定理3　$K[x]$ 中次数大于零的多项式 $f(x)$ 没有重因式的充分必要条件是 $f(x)$ 与它的导数 $f'(x)$ 互素.　□

定理3表明，判断一个多项式 $f(x)$ 有没有重因式，只要去计算 $(f(x),f'(x))$，而求最大公因式有统一的方法：辗转相除法，所以我们有统一的方法——辗转相除法判断一个多项式有没有重因式.不仅如此，由于在数域扩大时，两个多项式的互素性不改变，一个多项式的导数也不改变，因此我们还有下述结论：

命题4　设数域 F 包含数域 K，对于 $f(x)\in K[x]$，$f(x)$ 在 $K[x]$ 中没有重因式的充分必要条件是把 $f(x)$ 看成 $F[x]$ 中的多项式时，$f(x)$ 在 $F[x]$ 中没有重因式.即，$f(x)$ 有无重因式不会随数域的扩大而改变.　□

在一些问题中，如果多项式 $f(x)$ 有重因式，我们希望求出一个多项式 $g(x)$，它没有重因式，并且不计重数时，它与 $f(x)$ 含有完全相同的不可约因式.从下面的讨论中可得到 $g(x)$ 的求法.

设 $K[x]$ 中的多项式 $f(x)$ 的标准分解式是

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_m^{r_m}(x),$$

据定理 1 得

$$f'(x)=p_1^{r_1-1}(x)p_2^{r_2-1}(x)\cdots p_m^{r_m-1}(x)h(x),$$

其中 $h(x)$ 不能被任何 $p_i(x)$ 整除, $i=1,\cdots,m$. 于是

$$(f(x),f'(x))=p_1^{r_1-1}(x)p_2^{r_2-1}(x)\cdots p_m^{r_m-1}(x).$$

因此用 $(f(x),f'(x))$ 除 $f(x)$ 所得商式是

$$cp_1(x)p_2(x)\cdots p_m(x),$$

把它记作 $g(x)$,我们便得到一个没有重因式的多项式 $g(x)$,它与 $f(x)$ 含有完全相同的不可约因式(不计重数).

上面一段给出了去掉 $f(x)$ 的不可约因式重数的方法:先用辗转相除法求出 $(f(x),f'(x))$,然后对 $f(x)$ 与 $(f(x),f'(x))$ 做带余除法,所得商式 $g(x)$ 即为所求的没有重因式的多项式.

去掉 $f(x)$ 的不可约因式的重数有不少好处. 例如,为了求 $f(x)$ 的所有不可约因式,我们可以先用上述方法得到一个没有重因式的多项式 $g(x)$,它与 $f(x)$ 含有完全相同的不可约因式(不计重数),但由于 $g(x)$ 的次数小于 $f(x)$ 的次数,所以 $g(x)$ 的不可约因式可能比较容易求得. 如果我们求出了 $g(x)$ 的一个不可约因式 $p_i(x)$,那么用带余除法可求出 $p_i(x)$ 在 $f(x)$ 中的重数. 又如,在实际问题中常常要求出一个多项式 $f(x)$ 的根(关于多项式的根,我们在下节将详细讨论. 实际上,读者在中学代数里已知道多项式的根的概念),由于有些求多项式的根的算法只对没有重因式的多项式适用,因此我们可以先去掉 $f(x)$ 的不可约因式的重数,得到一个没有重因式的多项式 $g(x)$,而 $g(x)$ 与 $f(x)$ 有完全相同的根(不计重数).

例 1　证明: $\mathbb{Q}[x]$ 中的多项式

$$f(x)=1+x+\frac{x^2}{2!}+\cdots+\frac{x^n}{n!}$$

没有重因式.

证明　求 $f(x)$ 的导数得

$$f'(x)=1+x+\cdots+\frac{x^{n-1}}{(n-1)!},$$

于是 $f(x)=f'(x)+\dfrac{x^n}{n!}$. 据习题 7.3 的第 7 题的结果得

$$(f(x),f'(x))=\left(f'(x)+\frac{1}{n!}x^n,\ f'(x)\right)=\left(\frac{x^n}{n!},\ f'(x)\right).$$

由于 $\dfrac{x^n}{n!}$ 的不可约因式只有 x(不计重数),而 $x\nmid f'(x)$,所以

$$\frac{x^n}{n!}, f'(x) = 1,$$

从而$(f(x), f'(x)) = 1$.因此 $f(x)$没有重因式. x

习　题　7.5

1. 判别下列有理系数多项式有无重因式:

(1) $f(x) = x^3 - 3x^2 + 4$;

(2) $f(x) = x^3 + 2x^2 - 11x - 12$.

2. a, b应该满足什么条件,下述实系数多项式才能有重因式:

$$f(x) = x^3 + 2ax + b.$$

3. 举例说明: $K[x]$中,一个不可约多项式 $p(x)$是 $f(x)$的导数 $f'(x)$的 $k-1$重因式($k\geq 2$),但是 $p(x)$不是 $f(x)$的 k重因式.

4. 证明: $K[x]$中,若不可约多项式 $p(x)$是 $f(x)$的导数 $f'(x)$的 $k-1$重因式($k\geq 1$),并且 $p(x)$是 $f(x)$的因式,则 $p(x)$是 $f(x)$的 k重因式.

5. 证明: $K[x]$中,不可约多项式 $p(x)$是 $f(x)$的 k重因式($k\geq 1$)的充分必要条件为: $p(x)$是 $f(x), f'(x), \cdots, f^{(k-1)}(x)$的因式,但不是 $f^{(k)}(x)$的因式.

6. 在 $Q[x]$中求一个没有重因式的多项式 $g(x)$,使它与 $f(x)$含有完全相同的不可约因式(不计重数);然后求 $f(x)$的标准分解式:

$$f(x) = x^5 - 3x^4 + 2x^3 + 2x^2 - 3x + 1.$$

*7. 证明: $K[x]$中一个 n次($n\geq 1$)多项式 $f(x)$能被它的导数整除的充分必要条件是它与一个一次因式的 n次幂相伴.

§6　多项式的根,复数域上的不可约多项式

从唯一因式分解定理知道, $K[x]$中每一个次数大于零的多项式都能唯一地分解成数域 K上有限多个不可约多项式的乘积.由此看出,不可约多项式起着基本建筑块的作用,这促使我们去搞清楚 $K[x]$中不可约多项式有哪些?我们已经知道, $K[x]$中每一个一次多项式都是不可约的.于是需要进一步研究的是, $K[x]$中有没有次数大于1的不可约多项式?显然,在 $K[x]$中,如果 $p(x)$是次数大于1的不可约多项式,则 $p(x)$没有一次因式.从这点受到启发,首先需要研究 $K[x]$中一个多项式 $f(x)$有一次因式的充分必要条件.为此需要用一次多项式去除 $f(x)$,观察它的余式.

定理1(余数定理)　在 $K[x]$中,用 $x-a$去除 $f(x)$所得的余式是 $f(a)$.

证明　作带余除法,得

$$f(x)=h(x)(x-a)+r(x),\quad \deg r(x)<\deg(x-a).$$

若 $r(x)\neq 0$,则 $\deg r(x)=0$,记 $r(x)=r$,其中 $r\in K^*$.若 $r(x)=0$,则可以把 $r(x)$与 K 中的数 0 等同.综上所述有

$$f(x)=h(x)(x-a)+r,\qquad r\in K. \tag{1}$$

x 用 a 代入,从(1)式得, $f(a)=r$.因此用 $x-a$ 去除 $f(x)$所得的余式是 $f(a)$.

x

推论 2　在 $K[x]$中, $x-a$ 整除 $f(x)$当且仅当

$$f(a)=0.$$

x

从推论 2 以及其他许多问题受到启发,引出多项式的根的概念.

定义 1　设 K 是一个数域,设 R 是一个有单位元的交换环,且 R 可看成是 K 的一个扩环.对于 $K[x]$中一个多项式 $f(x)$,如果 R 中有一个元素 c 使得 $f(c)=0$,则称 c 是 $f(x)$在 R 中的一个根.

$f(x)$在复数域中的根称为复根.类似地,实系数多项式在实数域中的根称为实根;有理系数多项式在有理数域中的根称为有理根.

从定义 1 和推论 2 立即得到下述重要结论:

定理 3(Bezout 定理)　$K[x]$中, $x-a$ 整除 $f(x)$当且仅当 a 是 $f(x)$在 K 中的一个根.

x

从定理 3 看出, $K[x]$中的多项式 $f(x)$有一次因式的充分必要条件是 $f(x)$在 K 中有根.

利用根与一次因式的关系,对于 $K[x]$中的多项式在 K 中的根,我们可以定义重根的概念: $a\in K$ 称为 $f(x)\in K[x]$的一个 k 重根,如果$(x-a)$是 $f(x)$的 k 重因式.当 $k=1$ 时, a 称为单根;当 $k>1$ 时, a 称为重根.

利用根与一次因式的关系,我们可以得到 $K[x]$中的多项式 $f(x)$在 K 中的根的数目的一个上界:

定理 4　$K[x]$中的 n 次($n\geq 0$)多项式在 K 中至多有 n 个根(重根按重数计算).

证明　零次多项式没有根,因此结论成立.

设 $f(x)$是一个 n 次多项式, $n>0$.把 $f(x)$分解成不可约多项式的乘积.据定理 3 以及根的重数的定义容易看出, $f(x)$在 K 中的根的数目(重根按重数计算)等于分解式中一次因式的数目(重因式按重数计算),这个数目当然不超过 $f(x)$的次数 n.

x

从定理 4 可以得到一个重要推论:

推论 5　设 $K[x]$中两个多项式 $f(x)$与 $g(x)$的次数都不超过 n.如果 K 中有 $n+1$ 个不同元素 $a_1,a_2,\cdots,a_{n+1}$,使得 $f(a_i)=g(a_i)$, $i=1,2,\cdots,n+1$,

则 $f(x)=g(x)$.

证明　设 $h(x)=f(x)-g(x)$,假如 $h(x)\neq 0$,则

$$\deg h\leq \max\{\deg f,\deg g\}\leq n.$$

因为

$$h(a_i)=f(a_i)-g(a_i)=0,\quad i=1,2,\cdots,n+1,$$

所以 $h(x)$在 K 中至少有 $n+1$ 个不同的根.这与定理4矛盾.因此 $h(x)=0$.于是 $f(x)=g(x)$.　x

我们将在以后几处指出推论5的用处.

为了研究多项式的根,我们需要多项式函数的概念,并且研究多项式函数与多项式之间的关系.

设 $f(x)\in K[x]$,对于 K 中每一个元素 a,x 用 a 代入得,$f(a)\in K$.于是 $K[x]$中一个多项式 $f(x)$确定了 K 到 K 的一个映射(即 K 上的一个函数),仍用 f 表示,即

$$\begin{aligned} f:\ K &\quad K \\ a &\longmapsto f(a). \end{aligned}$$

这种由 $K[x]$中的多项式确定的 K 上的函数称为 K 上的多项式函数.

当 K 是实数域 R 时,上述定义的多项式函数就是数学分析中所讨论的多项式函数.

我们已经知道,$K[x]$中的每一个多项式都确定一个 K 上的多项式函数.现在要问:$K[x]$中两个不相等的多项式 $f(x)$与 $g(x)$,它们确定的 K 上的多项式函数 f 与 g 是否不相等?回答是肯定的,即我们有

定理6　数域 K 上的两个多项式 $f(x)$与 $g(x)$如果不相等,则它们确定的 K 上的多项式函数 f 与 g 也不相等.

证明　证逆否命题.设 K 上的多项式函数 f 与 g 相等,则对一切 $a\in K$,有 $f(a)=g(a)$.由于 K 是数域,它有无穷多个元素,于是据推论5得,$f(x)=g(x)$.即多项式 $f(x)$与 $g(x)$相等.　x

注意,定理6证明中的关键是 K 有无穷多个元素.

我们把数域 K 上的多项式函数组成的集合记作 K_{pol}.让 $K[x]$中的多项式 $f(x)$对应到它确定的 K 上的函数 f,这是 $K[x]$到 K_{pol}的一个映射,记作 σ.显然它是满射.定理6表明这个映射是单射,因此它是双射.

在集合 K_{pol}中规定加法与乘法如下:设 $f,g\in K_{pol}$,规定

$$(f+g)(a)\overset{\text{def}}{=}f(a)+g(a),\ \forall a\in K,$$

$$(fg)(a)\overset{\text{def}}{=}f(a)g(a),\ \forall a\in K.$$

容易验证，$f+g, fg \in K_{pol}$（设 f, g 分别是多项式 $f(x), g(x)$ 确定的函数，记

$$h(x)=f(x)+g(x),$$

则对任意 $a \in K$，有

$$h(a)=f(a)+g(a)=(f+g)(a),$$

因此 $h=f+g$. 这表明 $f+g$ 是多项式 $f(x)+g(x)$ 确定的函数. 类似可证 fg 是多项式 $f(x)g(x)$ 确定的函数)，并且 K_{pol} 是一个环，称它为 K 上的多项式函数环. 它的零元素是零多项式确定的函数 0，称为零函数；它的单位元素是多项式 1 确定的函数 1，它是一个常值函数，它把 K 中每个元素 a 映成 1. 容易看出 K_{pol} 是一个交换环.

环 $K[x]$ 与环 K_{pol} 有密切关系：

定理 7　设 K 是数域，让 $K[x]$ 中的多项式 $f(x)$ 对应到它确定的 K 上的函数 f，则这个映射 σ 是环 $K[x]$ 到环 K_{pol} 的双射，并且 σ 保持加法和乘法运算.

证明　前面已证 σ 是双射. 现在证 σ 保持运算.

设 $f(x), g(x) \in K[x]$，前面已证 $f(x)+g(x)$ 确定的函数是 $f+g$，$f(x)g(x)$ 确定的函数是 fg. 因此

$$\sigma(f(x)+g(x))=f+g=\sigma(f(x))+\sigma(g(x)),$$
$$\sigma(f(x)g(x))=fg=\sigma(f(x))\sigma(g(x)).$$

这证明了 σ 保持加法，并且保持乘法.　x

定义 2　设 R 和 R' 是两个环，如果存在从 R 到 R' 的一个双射 σ，它保持加法与乘法运算，即对任意 $a, b \in R$，有

$$\sigma(a+b)=\sigma(a)+\sigma(b),$$
$$\sigma(ab)=\sigma(a)\sigma(b),$$

那么称 σ 是环 R 到 R' 的一个同构映射，此时称环 R 与 R' 是同构的，记作 Rü R'.

定理 7 表明，数域 K 上的一元多项式环 $K[x]$ 与数域 K 上的多项式函数环 K_{pol} 是同构的. 同构的环，它们的元素之间有一个一一对应，并且这个对应保持加法与乘法运算. 因此在研究有关加法与乘法运算的问题时，同构的环具有相同的性质，从而可以把同构的环等同看待.

现在我们可以用多项式函数的概念来刻画 $K[x]$ 中的多项式在 K 中的根：$a \in K$ 是 $K[x]$ 中多项式 $f(x)$ 在 K 中的根当且仅当 K 上的多项式函数 f 在 a 处的函数值 $f(a)=0$. 于是我们可以运用函数论的知识来研究 $K[x]$ 中的多项式在 K 中的根，特别是当 K 为实数域 R 或复数域 C 的情形. 譬如，实系数多项式 $f(x)$ 的实根就是多项式函数 $y=f(x)$ 的图像与 x 轴的交点的横坐标. 据数学分析的知识，多项式函数是连续函数，而连续函数具有这样的性质：在区间

$[a,b]$上,它取 $f(a)$与 $f(b)$之间的每一个值.因此如果 $f(a)<0$,而 $f(b)>0$,则必有 $c\in(a,b)$,使得 $f(c)=0$,也就是说,在区间(a,b)上一定有多项式$f(x)$的一个根.

运用复变函数论的知识可以证明下面重要的定理.

定理8(代数基本定理)　每个次数大于零的复系数多项式在复数域中至少有一个根.

证明代数基本定理所用到的复变函数论的知识主要是下面的定理.

*定理9(Liouville定理)　若函数 $f(z)$的复平面 C 上解析,并且有界,则 $f(z)$必为一常值函数. x

* Liouville定理的证明可以在复变函数的教材中找到.这里我们只指出,函数 $f(z)$在复平面 C 上解析的意思是: $f(z)$在复平面 C 上的每一点的导数都存在.复变量函数的导数定义以及求导公式都与数学分析中实变量函数一样.例如,若函数 $f(z)$, $g(z)$在复平面 C 上解析,并且对于 C 上每一点 z,有 $g(z)\neq0$,则 $f(z)$ $g(z)$在复平面 C 上解析,并且

$$\left(\frac{f(z)}{g(z)}\right)'=\frac{f'(z)g(z)-f(z)g'(z)}{g^2(z)}.$$

*代数基本定理的证明　假如一个 n次$(n>0)$复系数多项式 $f(x)$在复数域中没有根,则对于复平面 C 上每一点 z,有 $f(z)\neq0$.由于多项式函数在复平面 C 上是解析的,于是函数

$$\varphi(z)=\frac{1}{f(z)}$$

在复平面上解析.由于 $z\to\infty$时,$\varphi(z)\to0$,所以存在正实数 R 和 M_1,使得当 $|z|>R$ 时,$|\varphi(z)|\leq M_1$.而当$|z|\leq R$ 时,连续函数 $\varphi(z)$在$|z|\leq R$ 上有界 M_2.因此对一切 $z\in C$,有

$$|\varphi(z)|\leq\max\{M_1,M_2\},$$

据 Liouville定理得,$\varphi(z)$必为一常值函数,从而 $f(z)$为一个常值函数,这与多项式 $f(x)$的次数大于零矛盾. x

定理8被称为"代数基本定理"是因为在19世纪以前解代数方程是代数学的最重要课题.在现代,定理8被看成代数学的基本定理之一(但不是最基本的).这个定理的第一个严格证明是高斯(Gauss)在1799年给出的,后来他又给出了四个证明.Jordan,Weyl等人也给过证明.

利用根与一次因式的关系(即Bezout定理),代数基本定理可以等价地叙述为:

定理10　每一个次数大于零的复系数多项式,在复数域上一定有一个一次因式.

由此可知,复数域上所有次数大于1的多项式全是可约的.换句话说,复数

域上的不可约多项式只有一次多项式.于是唯一因式分解定理在复数域上可以叙述成:

定理 11(复系数多项式唯一因式分解定理) 每个次数大于零的复系数多项式在复数域上都可以唯一地分解成一次因式的乘积. x

因此,复系数多项式 $f(x)$具有标准分解式

$$f(x)=a(x-c_1)^{r_1}(x-c_2)^{r_2}\cdots(x-c_m)^{r_m}, \tag{2}$$

其中 a 是 $f(x)$的首项系数,$c_1,c_2,\cdots,c_m$ 是不同的复数,$r_1,r_2,\cdots,r_m$ 是正整数,从标准分解式(2)得出

推论 12 每个 n 次($n\geq 0$)复系数多项式恰有 n 个复根(重根按重数计算). x

现在我们用复系数多项式唯一因式分解定理来讨论数域 K 上的多项式 $f(x)$的复根与 $f(x)$的系数之间的关系.

设 $f(x)$是 $K[x]$中首项系数为 1 的 n 次($n\geq 1$)多项式,把 $f(x)$看成复系数多项式,则 $f(x)$有 n 个复根 $c_1,c_2,\cdots,c_n$(这里可能有相同的),于是在复数域上 $f(x)$有因式分解

$$f(x)=(x-c_1)(x-c_2)\cdots(x-c_n). \tag{3}$$

设 $f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,\ a_i\in K,\ i=0,1,\cdots,n-1,$

把(3)式的右端展开,并且比较(3)式两端的多项式的系数便得到

$$a_{n-1}=-(c_1+c_2+\cdots+c_n),$$

$$\cdots\cdots\cdots\cdots$$

$$a_{n-k}=(-1)^k\sum_{1\leq i_1<i_2<\cdots<i_k\leq n}c_{i_1}c_{i_2}\cdots c_{i_k}, \tag{4}$$

$$\cdots\cdots\cdots\cdots$$

$$a_0=(-1)^n c_1c_2\cdots c_n.$$

这些公式是用多项式 $f(x)$的复根表示它的系数的表达式,称为 Vieta 公式.注意公式中的 $c_1,c_2,\cdots,c_n$ 是复数,但是 $a_{n-1},a_{n-2},\cdots,a_0$ 是数域 K 中的元素.

如果 $f(x)$首项系数 $a_n\neq 1$,则公式(4)给出的是比值$\dfrac{a_i}{a_n}$的表达式.

习 题 7.6

1. 设 $f(x)=x^5+7x^4+19x^3+26x^2+20x+8\in Q[x]$,判断 -2 是不是多项式 $f(x)$的根,如果是的话,它是几重根?

(提示:用综合除法看 $x+2$ 能否整除 $f(x)$.)

2. 设 $f(x)=2x^3-7x^2+4x+a\in Q[x]$,求 a 值,使 $f(x)$ 有重根,并且求出相应的重根及其重数.

(提示: $c\in Q$ 是 $f(x)$ 的重根　　$x-c$ 是 $f(x)$ 的重因式　　$x-c$ 是 $f(x)$ 与 $f'(x)$ 的公因式　　$x-c$ 是 $(f(x),f'(x))$ 的因式.)

3. 求下列复系数多项式在复数域中的公共根:

$f(x)=x^3-x^2-x-2$;　$g(x)=x^4-2x^3+2x^2-3x-2$.

(提示: $a\in C$ 是 $f(x)$ 与 $g(x)$ 公共根　　$x-a$ 是 $f(x)$ 与 $g(x)$ 的公因式　　$x-a$ 是 $(f(x),g(x))$ 的因式.)

4. 设 $f(x)=x^4-5x^3+ax^2+bx+9\in Q[x]$,如果 3 是 $f(x)$ 的二重根,求 a,b.

5. 证明:在 $K[x]$ 中,如果 $(x+1)\mid f(x^{2n+1})$,则 $(x^{2n+1}+1)\mid f(x^{2n+1})$.

6. 设 $f_1(x),f_2(x)\in Q[x]$,证明:如果

$$(x^2+x+1)\mid f_1(x^3)+xf_2(x^3),$$

则 1 是 $f_i(x)$ 的根, $i=1,2$.

7. 设 $C[x]$ 中的多项式 $f(x)\neq 0$,并且 $f(x)\mid f(x^m)$,其中 m 是一个大于 1 的整数.证明 $f(x)$ 在 C 中的根只能是零或单位根.

8. 求复系数多项式 x^n-1 的标准分解式.

*9. 设 K 是一个数域,设 R 是(或者可看成是) K 的一个交换扩环.设 $a\in R$,且 $J_a\neq\{0\}$,其中 J_a 表示 $K[x]$ 中那些多项式组成的集合,它们中每一个在 R 中有根 a,即

$$J_a=\{f(x)\in K[x]\mid f(a)=0\}.$$

证明:(1) J_a 中存在唯一的首项系数为 1 的多项式 $m(x)$,使得 J_a 的每个多项式都是 $m(x)$ 的倍式;

(2) 如果 R 是无零因子环,则 $m(x)$ 在 $K[x]$ 中不可约.

10. 证明:数域 K 上任意一个不可约多项式在复数域内没有重根.

11. 设 A 是数域 K 上的 n 级矩阵,证明: A 的特征多项式的 n 个复根的和等于 $\mathrm{tr}(A)$, n 个复根的乘积等于 $|A|$.

12. 设 $p(x)$ 是数域 K 上的一个不可约多项式,证明:如果 $K[x]$ 中的多项式 $g(x)$ 与 $p(x)$ 有一个公共复根,则在 $K[x]$ 中 $p(x)\mid g(x)$.

§7　实数域上的不可约多项式

这一节我们要找出实数域上的所有不可约多项式.由于每一个复数都可以表示成 $a+bi$ 的形式,其中 a,b 是实数,因此有可能利用复数域上多项式的信息来研究实数域上的不可约多项式.

定理 1　设 $f(x)$ 是实系数多项式,如果 c 是 $f(x)$ 的一个复根,则 c 的共轭复数珋也是 $f(x)$ 的一个复根.

证明　设

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0,\ a_i \in R,\ i = 0, 1, \cdots, n,$$

因为 c 是 $f(x)$ 的复根,所以

$$f(c) = a_n c^n + a_{n-1} c^{n-1} + \cdots + a_1 c + a_0 = 0. \tag{1}$$

在(1)式两边取共轭复数得

$$a_n \bar{c}^n + a_{n-1} \bar{c}^{n-1} + \cdots + a_1 \bar{c} + a_0 = 0, \tag{2}$$

即 $f(\bar{c}) = 0$.因此 $\bar{c}$ 是 $f(x)$ 的一个复根. □

利用定理 1 我们可以得出

定理 2　实数域上的不可约多项式都是一次多项式或判别式小于零的二次多项式.

证明　设 $f(x) \in R[x]$ 是不可约的,把 $f(x)$ 看成复系数多项式,根据代数基本定理,$f(x)$ 有一个复根 c.如果 c 是实数,则 $f(x)$ 在 $R[x]$ 中有一次因式 $x - c$.因为 $f(x)$ 不可约,所以

$$(x - c) \sim f(x),$$

即 $f(x) = a(x - c)$,a 是非零实数.因此 $f(x)$ 是一次多项式.

如果 c 是虚数,据定理 1,$\bar{c}$ 也是 $f(x)$ 的一个复根.由于 $\bar{c} \neq c$,所以 $(x - c, x - \bar{c}) = 1$.在 $C[x]$ 中

$$(x - c) \mid f(x), \quad (x - \bar{c}) \mid f(x),$$

从而据互素多项式的性质 2 得

$$(x - c)(x - \bar{c}) \mid f(x).$$

我们有

$$(x - c)(x - \bar{c}) = x^2 - (c + \bar{c})x + c\bar{c},$$

而 $c + \bar{c}$ 与 $c\bar{c}$ 都是实数.既然在 $C[x]$ 中

$$(x^2 - (c + \bar{c})x + c\bar{c}) \mid f(x),$$

于是在 $R[x]$ 中

$$(x^2 - (c + \bar{c})x + c\bar{c}) \mid f(x).$$

由于 $f(x)$ 在 $R[x]$ 中不可约,因此

$$f(x) = a(x^2 - (c + \bar{c})x + c\bar{c}),$$

a 是非零实数.此时 $f(x)$ 是二次多项式.

实系数二次多项式 $f(x) = ax^2 + bx + c$ 不可约

$f(x)$ 在 $R[x]$ 中没有一次因式

$f(x)$ 没有实根

$b^2 - 4ac < 0$. □

从定理 2 和唯一因式分解定理立即得出

定理 3(实系数多项式唯一因式分解定理)　每个次数大于零的实系数多项式在实数域上都可以唯一地分解成一次因式与判别式小于零的二次因式的乘

积.　　x

因此,实系数多项式具有标准分解式

$$f(x) = a(x - c_1)^{r_1} \cdots (x - c_s)^{r_s} \cdot (x^2 + p_1 x + q_1)^{k_1} \cdots (x^2 + p_t x + q_t)^{k_t}, \tag{3}$$

其中 a 是 $f(x)$的首项系数;$c_1, \cdots, c_s$ 是不同的实数;$p_1, q_1; p_2, q_2; \cdots; p_t, q_t$ 是不同的实数对,并且满足 $p_i^2 - 4q_i < 0$, $i = 1, \cdots, t$; $r_1, \cdots, r_s$, $k_1, \cdots, k_t$ 都是非负整数.

从实系数多项式 $f(x)$的标准分解式(3)看出,如果虚数 b 是 $f(x)$的一个复根,则 b 的共轭复数$\bar{b}$也是 $f(x)$的一个复根,并且它们的重数相同(此时把 $f(x)$看成复系数多项式).因此通常我们说,实系数多项式的虚根共轭成对出现.

习　题　7.7

1. 证明:实系数的奇次多项式至少有一个实根.

2. 求多项式 $x^n - 1$ 在实数域上的标准分解式.

3. 设 A 是实数域上的 n 级斜对称矩阵,它的特征多项式$|\lambda I - A|$记作 $f(\lambda)$.证明:(1) $f(\lambda)$的复根都是纯虚数或零;

(2) 如果 A 可逆,则 $f(\lambda)$的不可约因式都是二次的.

*4. 证明:如果 n 级实矩阵 A 与 B 不相似,则把它们看成复矩阵后仍然不相似.

§8　有理数域上的不可约多项式

这一节讨论有理数域上的不可约多项式有哪些?如何判别一个有理系数多项式是否不可约?这些问题的回答比复系数多项式和实系数多项式困难得多.

在本章§6 的开头曾指出,在 $K[x]$中,如果一个次数大于 1 的多项式 $p(x)$不可约,则 $p(x)$没有一次因式,从而 $p(x)$在 K 中没有根,这样就可以缩小讨论 $Q[x]$中不可约多项式的范围.那么如何判别 $Q[x]$中次数大于 1 的多项式 $f(x)$有没有有理根呢?显然 $f(x)$有有理根当且仅当 $f(x)$在 $Q[x]$中的相伴元也有有理根.因此很自然地选取 $f(x)$在 $Q[x]$中的一个最简单的相伴元来研究.让我们先看一个例子.

$$f(x) = \frac{1}{2}x^4 + \frac{1}{3}x^3 - 2x + 1$$

$$=\frac{1}{6}(3x^4+2x^3-12x+6),$$

记 $g(x)=3x^4+2x^3-12x+6$,则 $f(x)\sim g(x)$.注意 $g(x)$是整系数多项式,而且它的各项系数的最大公因数只有±1.由此受到启发,引出下述概念:

定义 1 一个非零的整系数多项式

$$g(x)=b_nx^n+\cdots+b_1x+b_0,$$

如果它的各项系数的最大公因数只有±1,则称 $g(x)$是一个本原多项式.

从上面的例子看出,任何一个非零的有理系数多项式 $f(x)$都与一个本原多项式 $g(x)$相伴.事实上,这只要先求出 $f(x)$的各项系数的分母的最小公倍数 m,然后提出$\frac{1}{m}$;再将括号内的整系数多项式的各项系数的公因数提出,这时括号内的整系数多项式便是本原多项式.

对于一个非零的有理系数多项式 $f(x)$,它可以与几个本原多项式相伴?下面的引理 1 就可以回答这个问题.

引理 1 两个本原多项式 $g(x)$与 $h(x)$在 $Q[x]$中相伴当且仅当 $g(x)=\pm h(x)$.

证明 充分性是显然的.下面证必要性.

设 $g(x)$与 $h(x)$在 $Q[x]$中相伴,则存在 $r\in Q$,使得 $g(x)=rh(x)$.设 $g(x)=\sum_{i=0}^{n}b_ix^i$, $h(x)=\sum_{i=0}^{n}c_ix^i$,其中 $b_i,c_i\in Z$, $i=0,1,\cdots,n$.假如 $r\neq\pm1$,设 $r=\frac{q}{p}$,其中$(p,q)=1$.于是 p,q 中至少有一个不等于±1,不妨说 $p\neq\pm1$.从而 $pg(x)=qh(x)$.比较 i 次项的系数得,$pb_i=qc_i$.于是 $p\mid qc_i$.由于$(p,q)=1$,因此 $p\mid c_i$, $i=0,1,2,\cdots,n$,这与 $h(x)$是本原多项式矛盾.因此 $r=\pm1$,即 $g(x)=\pm h(x)$. x

引理 1 告诉我们,对于一个非零的有理系数多项式 $f(x)$,与它在 $Q[x]$中相伴的本原多项式有且只有两个,它们相差一个正负号.现在我们来研究本原多项式在 $Q[x]$中的不可约性问题.为此首先介绍本原多项式的一个重要性质,即

引理 2(高斯(Gauss)引理) 两个本原多项式的乘积还是本原多项式.

证明 设

$$f(x)=a_nx^n+\cdots+a_1x+a_0,$$

$$g(x)=b_mx^m+\cdots+b_1x+b_0$$

是两个本原多项式.设

$$h(x)=f(x)g(x)=c_{n+m}x^{n+m}+\cdots+c_1x+c_0,$$

其中 $c_s = \sum_{i+j=s} a_i b_j, s=0,1,\cdots,n+m$.

假如 $h(x)$不是本原多项式,则存在一个素数 p,使得 $p \mid c_s, s=0,1,2,\cdots,n+m$.因为 $f(x)$是本原的,所以 p 不能同时整除 $f(x)$的每一项的系数.于是存在 $k(0\leq k\leq n)$满足

$$p \mid a_0,\ p \mid a_1,\cdots,\ p \mid a_{k-1},\ p \nmid a_k. \tag{1}$$

同理,存在 $l(0\leq l\leq m)$满足

$$p \mid b_0,\ p \mid b_1,\cdots,\ p \mid b_{l-1},\ p \nmid b_l. \tag{2}$$

考虑 $h(x)$的 $k+l$次项的系数:

$$\begin{aligned} c_{k+l} = {} & a_{k+l} b_0 + a_{k+l-1} b_1 + \cdots + a_{k+1} b_{l-1} \\ & + a_k b_l + a_{k-1} b_{l+1} + \cdots + a_0 b_{k+l}, \end{aligned} \tag{3}$$

由(1),(2)两式得,$p \nmid c_{k+l}$(注意从 $p \nmid a_k$ 且 $p \nmid b_l$ 可推出 $p \nmid a_k b_l$),矛盾.因此 $h(x)$是本原多项式. □

定理1　一个次数大于0的本原多项式 $g(x)$在 Q 上可约当且仅当 $g(x)$可以分解成两个次数都比 $g(x)$的次数低的本原多项式的乘积.

证明　充分性是显然的.下面证必要性.

设本原多项式 $g(x)$在 Q 上可约,则存在 $g_i(x)\in Q[x], i=1,2$,使得

$$g(x) = g_1(x) g_2(x), \quad \deg g_i(x) < \deg g(x),\ i=1,2.$$

设 $g_i(x) = r_i h_i(x)$,其中 $r_i \in Q$, $h_i(x)$是本原多项式, $i=1,2$,则

$$g(x) = r_1 r_2 h_1(x) h_2(x).$$

由于 $h_1(x) h_2(x)$也是本原多项式,因此 $r_1 r_2 = \pm 1$.从而 $g(x) = [\pm h_1(x)] h_2(x)$.由于 $\deg(\pm h_1(x)) = \deg g_1(x) < \deg g(x)$, $\deg h_2(x) = \deg g_2(x) < \deg g(x)$,因此 $g(x)$分解成了两个次数较低的本原多项式的乘积. □

推论2　如果一个次数大于0的整系数多项式 $f(x)$在 Q 上可约,则它可以分解成两个次数比它低的整系数多项式的乘积.

证明　设 $f(x) = rg(x)$,其中 $g(x)$是本原多项式, $r\in Z$,据定理1得, $g(x) = h_1(x) h_2(x)$,其中 $h_i(x)$是本原多项式,且 $\deg h_i(x) < \deg g(x)$, $i=1,2$.从而 $f(x) = [rh_1(x)] h_2(x)$.这表明 $f(x)$分解成了两个次数较低的整系数多项式的乘积. □

定理3　每一个次数大于0的本原多项式 $g(x)$可以唯一地分解成 Q 上不可约的本原多项式的乘积.唯一性是指,假如 $g(x)$有两个这样的分解式:

$$g(x) = p_1(x) p_2(x) \cdots p_s(x), \quad g(x) = q_1(x) q_2(x) \cdots q_t(x),$$

则 $s=t$,且适当排列因式的次序后,有

$$p_i(x) = \pm q_i(x), \quad i=1,2,\cdots,s.$$

证明 可分解性.对本原多项式的次数 n 作数学归纳法. $n=1$ 时,$g(x)=g(x)$,且 $g(x)$不可约.假设对于次数小于 n 的本原多项式可分解性成立.现在来看 n 次本原多项式 $g(x)$.如果 $g(x)$在 Q 上不可约,则 $g(x)=g(x)$.如果 $g(x)$在 Q 上可约,则据定理 1 得,$g(x)=g_1(x)g_2(x)$,其中 $g_i(x)$是本原多项式,且 $\deg g_i(x)<\deg g(x)$, $i=1,2$.对 $g_i(x)$用归纳假设便得出 $g(x)$的可分解性.

唯一性.据 $Q[x]$中唯一因式分解定理得,$s=t$,且经过适当排列因式的次序后,有 $p_i(x)\sim q_i(x)$, $i=1,2,\cdots,s$.由于 $p_i(x)$, $q_i(x)$都是本原多项式,因此 $p_i(x)=\pm q_i(x)$, $i=1,2,\cdots,s$. x

利用定理 3 可以得到求整系数多项式的全部有理根的方法:

定理 4 设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

是一个次数 n 大于 0 的整系数多项式,如果$\frac{q}{p}$是 $f(x)$的一个有理根,其中 p, q 是互素的整数,那么

$$p\mid a_n,\quad q\mid a_0.$$

证明 设 $f(x)=rf_1(x)$,其中 $r\in Z^*$, $f_1(x)$是本原多项式,设$\frac{q}{p}$是 $f(x)$的一个根,则$\frac{q}{p}$也是 $f_1(x)$的一个根.于是在 $Q[x]$中,$\left(x-\frac{q}{p}\right)\Big| f_1(x)$,从而 $(px-q)\mid f_1(x)$.由于$(p,q)=1$,因此 $px-q$ 是本原多项式.据定理 3 和高斯引理得

$$f_1(x)=(px-q)g(x),$$

其中 $g(x)=b_{n-1}x^{n-1}+\cdots+b_1x+b_0$ 是本原多项式.于是

$$f(x)=r(px-q)g(x). \tag{4}$$

分别比较(4)式两边多项式的首项系数与常数项得

$$a_n=rpb_{n-1},\quad a_0=-rqb_0.$$

因此

$$p\mid a_n,\quad q\mid a_0.$$ x

定理 4 的证明过程表明,如果$\frac{q}{p}$是 $f(x)$的一个有理根,且$(p,q)=1$,则存在一个整系数多项式 $g(x)$使得

$$f(x)=(px-q)g(x). \tag{5}$$

从(5)式得

$$f(1)=(p-q)g(1),\quad f(-1)=-(p+q)g(-1).$$

当$\frac{q}{p}\neq\pm1$时,有

$$\frac{f(1)}{p-q}=g(1)\in Z,\quad \frac{f(-1)}{p+q}=-g(-1)\in Z.$$

因此如果计算出$\frac{f(1)}{p-q}|\ Z$ 或$\frac{f(-1)}{p+q}|\ Z$,则可断言$\frac{q}{p}$不是 $f(x)$的根.这个判断方法在求整系数多项式的有理根时有用.

例 1　判断 $f(x)=x^3+x+1$ 在有理数域上是否不可约?

解　如果 $f(x)$在 Q 上可约,则 $f(x)$可以分解成两个次数较低的有理系数多项式的乘积.由于 $\deg f(x)=3$,因此 $f(x)$在 $Q[x]$中必有一次因式,从而 $f(x)$必有有理根.由于 $a_3=1$, $a_0=1$,因此据定理 4 得, $f(x)$的有理根只可能是 ±1.但是

$$f(1)=3,\quad f(-1)=-1.$$

因此±1 都不是 $f(x)$的根.这个矛盾表明 $f(x)$在 Q 上不可约.

注意:如果整系数多项式 $f(x)$的次数大于 3,那么不能从 $f(x)$没有有理根便得出 $f(x)$在 Q 上不可约的结论.这是因为 $f(x)$没有有理根,只是说明 $f(x)$没有一次因式.但是 $f(x)$可能有次数大于 1 的因式,从而 $f(x)$可能可约.

例 2　求 $f(x)=3x^4+8x^3+6x^2+3x-2$ 的全部有理根.

解　$a_4=3$ 的因子只有±1, ±3; $a_0=-2$ 的因子只有±1, ±2.据定理 4 得, $f(x)$的有理根只可能是:

$$\pm1,\quad \pm2,\quad \pm\frac{1}{3},\quad \pm\frac{2}{3}.$$

因为 $f(1)=18\neq0$, $f(-1)=-4\neq0$,所以±1 不是 $f(x)$的根.

考虑 2,因为

$$\frac{f(-1)}{p+q}=-\frac{4}{3}$$

不是整数,所以 2 不是 $f(x)$的根.

考虑 -2,因为

$$\frac{f(1)}{p-q}=\frac{18}{3}=6,\quad \frac{f(-1)}{p+q}=\frac{-4}{-1}=4,$$

所以需要进一步用综合除法来判断 -2 是不是 $f(x)$的根:

$$\begin{array}{rrrrr|r}
3 & 8 & 6 & 3 & -2 & -2 \\
 & -6 & -4 & -4 & 2 & \\
\hline
3 & 2 & 2 & -1 & 0 & \\
 & -6 & 8 & -20 & & \\
\hline
3 & -4 & 10 & -21 & &
\end{array}$$

这表明 -2 是 $f(x)$的单根.于是

$$f(x)=(x+2)(3x^3+2x^2+2x-1).$$

考虑$\frac{1}{3}$,因为

$$\frac{f(1)}{p-q}=9,\quad \frac{f(-1)}{p+q}=-1,$$

所以需作综合除法:

$$\begin{array}{rrrr|l}
3 & 2 & 2 & -1 & \frac{1}{3} \\
 & 1 & 1 & 1 & \\
\hline
3 & 3 & 3 & 0 & \\
 & 1 & \frac{4}{3} & & \\
\hline
3 & 4 & \frac{13}{3} & &
\end{array}$$

这表明$\frac{1}{3}$是 $f(x)$的单根.于是

$$f(x)=(x+2)\left(x-\frac{1}{3}\right)(3x^2+3x+3).$$

显然 x^2+x+1 没有有理根.因此 $f(x)$的全部有理根是 $-2,\frac{1}{3}$,它们都是单根.

下面给出判别一个整系数多项式在有理数域上不可约的常用方法.

定理 5(Eisenstein 判别法)　设

$$f(x)=a_nx^n+\cdots+a_1x+a_0$$

是一个次数 n 大于零的整系数多项式,如果存在一个素数 p,使得

1) $p\nmid a_n$;

2) $p\mid a_i$, $i=0,1,\cdots,n-1$;

3) $p^2\nmid a_0$;

则 $f(x)$在有理数域上是不可约的.

证明　假如 $f(x)$在 Q 上可约,则据推论 2 得

$$f(x)=(b_mx^m+\cdots+b_1x+b_0)(c_lx^l+\cdots+c_1x+c_0),\tag{6}$$

其中 b_i, c_j 都是整数($i=0,1,\cdots,m$; $j=0,1,\cdots,l$), $b_m\neq 0$, $c_l\neq 0$, $m<n$, $l<n$,并且 $m+l=n$.因此

$$a_n=b_mc_l,\quad a_0=b_0c_0.$$

已知 $p\mid a_0$,所以 $p\mid b_0$ 或者 $p\mid c_0$.又因为 $p^2\nmid a_0$,所以 p 不能同时整除 b_0 和 c_0.不妨设 $p\mid b_0$ 但 $p\nmid c_0$.因为 $p\nmid a_n$,所以 $p\nmid b_m$.假设 $b_0,b_1,\cdots,b_m$ 中第一个不能

被 p 整除的是 b_k，即

$$p \mid b_0, p \mid b_1, \cdots, p \mid b_{k-1}, p \nmid b_k, 0 < k \leq m. \tag{7}$$

比较(6)式两边 x^k 的系数，得

$$a_k = b_k c_0 + b_{k-1} c_1 + \cdots + b_0 c_k, \tag{8}$$

因为 $k \leq m < n$，所以 $p \mid a_k$. 于是从(7)和(8)式得 $p \mid b_k c_0$. 由于 $p \nmid b_k$，从而 $p \mid c_0$，矛盾. 所以 $f(x)$ 在 Q 上不可约. □

利用定理 5 我们可以得到

推论 6 在 $Q[x]$ 中存在任意次数的不可约多项式.

证明 对任意的正整数 n，设 $f(x) = x^n + 2$. 素数 2 符合定理 5 的条件，因此 $x^n + 2$ 在 Q 上不可约. □

有时直接用 Eisenstein 判别法无法判断 $f(x)$ 是否在 Q 上不可约，我们可以通过不定元 x 用 $Q[x]$ 中的元素 $x + b$ 代入，得到另一个多项式

$$g(x) := f(x + b)$$

对 $g(x)$ 可能用 Eisenstein 判别法能判断它不可约，这时原来的 $f(x)$ 是否不可约？下面的命题回答了这个问题：

命题 7 设 $f(x) = a_n x^n + \cdots + a_1 x + a_0$ 是次数 n 大于零的整系数多项式，设 b 是任意给定的一个整数. 令

$$\begin{aligned} g(x) &:= f(x + b) \\ &= a_n (x + b)^n + \cdots + a_1 (x + b) + a_0, \end{aligned}$$

则 $f(x)$ 在 Q 上不可约的充分必要条件是 $g(x)$ 在 Q 上不可约.

证明 先证充分性. 因为 $g(x)$ 的首项为 $a_n x^n$，所以

$$\deg g(x) = n = \deg f(x).$$

假如 $f(x)$ 在 Q 上可约，则

$$f(x) = f_1(x) f_2(x), \quad \deg f_i < \deg f, \quad i = 1, 2. \tag{9}$$

x 用 $x + b$ 代入，从(9)式得

$$f(x + b) = f_1(x + b) f_2(x + b), \tag{10}$$

(10)式表明 $g(x)$ 在 Q 上可约.

再证必要性. x 用 $x - b$ 代入，则从

$$g(x) = f(x + b),$$

得

$$g(x - b) = f(x).$$

从上面证得的结论知，如果 $f(x)$ 在 Q 上不可约，则 $g(x)$ 在 Q 上也不可约. □

例 3 设 p 是一个素数，多项式

$$f(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$$

称为一个分圆多项式. 证明 $f(x)$ 在 Q 上不可约.

证明　我们有

$$(x-1)f(x)=(x-1)(x^{p-1}+x^{p-2}+\cdots+x+1)$$
$$=x^p-1. \tag{11}$$

x 用 $x+1$ 代入,从(11)式得

$$xf(x+1)=(x+1)^p-1. \tag{12}$$

从(12)式得

$$f(x+1)=x^{p-1}+px^{p-2}+\cdots+C_p^k x^{p-k-1}+\cdots+p.$$

令　$g(x):=f(x+1)$

$$=x^{p-1}+px^{p-2}+\cdots+C_p^k x^{p-k-1}+\cdots+p,$$

因为　$C_p^k=\dfrac{p(p-1)\cdots(p-k+1)}{k!}, 1\leq k<p,$

并且 $(p,k!)=1$,所以

$$k!\mid(p-1)\cdots(p-k+1),$$

从而　$$p\mid C_p^k, 1\leq k<p.$$

于是对于 $g(x)$,素数 p 满足 Eisenstein 判别法的条件,所以 $g(x)$ 在 Q 上不可约.从而 $f(x)$ 在 Q 上不可约.　x

对于有些整系数多项式,即使用命题 7 仍然不能用 Eisenstein 判别法判别其是否不可约.对于二次或三次整系数多项式 $f(x)$,如果它可约,则它一定有一次因式,从而它必有有理根.于是只要考察 $f(x)$ 是否有有理根,那么就可以解决二次或三次整系数多项式是否不可约的问题.对于大于 3 次的整系数多项式,则还需要考虑它是否有大于 1 次的因式.

习　题　7.8

1. 求下列多项式的全部有理根:

(1) $2x^3+x^2-3x+1$;

(2) $2x^4-x^3-19x^2+9x+9$.

2. 下列整系数多项式在有理数域上是否不可约?

(1) $x^4-6x^3+2x^2+10$;　(2) $7x^5+18x^4+6x-6$;

(3) x^5+5x^3+1;　(4) x^4-2x^3+2x-3;

(5) x^p+px^2+1, p 为奇素数;

(6) x^3+x^2-3x+2;

(7) $2x^3-x^2+x+1$;　(8) x^4-5x+1.

3. 设 $n>1$,证明: n 个互不相同的素数的几何平均数一定是无理数.

4. 设 $f(x)=a_nx^n+\cdots+a_1x+a_0$ 是一个次数大于零的整系数多项式,证明:如果

$a_n + a_{n-1} + \cdots + a_1 + a_0$ 是一个奇数,则 1 和 -1 都不是 $f(x)$的根.

5. 设 $f(x)$是一个次数大于零且首项系数为 1 的整系数多项式,证明:如果 $f(0)$与 $f(1)$都是奇数,则 $f(x)$没有有理根.

6. 设 $f(x) = x^3 + ax^2 + bx + c$ 是整系数多项式,证明:如果$(a + b)c$是奇数,则 $f(x)$在有理数域上不可约.

7. 设 $a_1, a_2, \cdots, a_n$ 是 n 个不同的整数,设

$$f(x) = (x - a_1)(x - a_2)\cdots(x - a_n) + 1,$$

证明:如果 n 是奇数,则 $f(x)$在有理数域上不可约.如果 n 是偶数, $f(x)$是否在有理数域上不可约?

*8. 设 $a_1, a_2, \cdots, a_n$ 是 n 个不同的整数,设

$$f(x) = (x - a_1)(x - a_2)\cdots(x - a_n) - 1,$$

证明 $f(x)$在有理数域上不可约.

*9. 设 $a_1, a_2, \cdots, a_n$ 是 n 个不同整数,设

$$f(x) = \prod_{i=1}^{n}(x - a_i)^2 + 1,$$

证明 $f(x)$在 Q 上不可约.

§9　多元多项式环

在本套教材的上册第 6 章中,我们讲了数域 K 上的 n 元二次型

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j,$$

其中 $a_{ij} \in K, 1 \leq i, j \leq n$.现在我们把它推广.

定义 1　设 K 是一个数域,用不属于 K 的 n 个符号 $x_1, x_2, \cdots, x_n$ 作表达式

$$\sum_{i_1, i_2, \cdots, i_n} a_{i_1 i_2 \cdots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}, \tag{1}$$

其中 $a_{i_1 i_2 \cdots i_n} \in K$, $i_1, i_2, \cdots, i_n$ 是非负整数,(1)式中的每一项称为一个单项式,(1)式称为数域 K 上的 n 元多项式,如果它具有下述性质:只有有限多个单项式的系数不为 0,并且两个这种形式的表达式相等当且仅当它们除去系数为零的单项式外含有完全相同的单项式,而系数为 0 的单项式允许任意删去和添进来.此时,符号 $x_1, x_2, \cdots, x_n$ 称为 n 个无关不定元.

在数域 K 上的 n 元多项式中,如果两个单项式的 $x_j (j = 1, 2, \cdots, n)$的幂指数都对应相等,则称这两个单项式为同类项.我们约定 n 元多项式中的单项式都是不同类的(即要把同类项合并成一项).

数域 K 上一个 n 元多项式如果它的所有系数全为 0,则称它为零多项式,记为 0.

我们常用 $f(x_1,x_2,\cdots,x_n)$,$g(x_1,x_2,\cdots,x_n)$,⋯或者 $f,g,\cdots$来代表 n 元多项式.

对于单项式 $a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$, $i_1+i_2+\cdots+i_n$ 称为它的全次数(或者简称为次数).

一个 n 元多项式 $f(x_1,x_2,\cdots,x_n)$的系数不为 0 的单项式的次数的最大值称为 $f(x_1,x_2,\cdots,x_n)$的全次数(或者简称为次数),用 $\deg f$ 来表示.

零多项式的全次数(简称为次数)规定为 $-\infty$.

例如,3 元多项式

$$5x_1^4+3x_1^3x_2+2x_1x_2x_3^2+x_2^3+x_2x_3$$

的次数是 4,其中单项式 $5x_1^4$,$3x_1^3x_2$,$2x_1x_2x_3^2$ 的次数都是 4.

数域 K 上所有 n 元多项式组成的集合记作 $K[x_1,x_2,\cdots,x_n]$. 在 $K[x_1,x_2,\cdots,x_n]$中定义加法与乘法如下:

$$\sum_{i_1,i_2,\cdots,i_n}a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}+\sum_{i_1,i_2,\cdots,i_n}b_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$$
$$\overset{\text{def}}{=}\sum_{i_1,i_2,\cdots,i_n}(a_{i_1i_2\cdots i_n}+b_{i_1i_2\cdots i_n})x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n} \tag{2}$$

$$\sum_{i_1,i_2,\cdots,i_n}a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}\sum_{j_1,j_2,\cdots,j_n}b_{j_1j_2\cdots j_n}x_1^{j_1}x_2^{j_2}\cdots x_n^{j_n}$$
$$\overset{\text{def}}{=}\sum_{s_1,s_2,\cdots,s_n}c_{s_1s_2\cdots s_n}x_1^{s_1}x_2^{s_2}\cdots x_n^{s_n}, \tag{3}$$

其中 $$c_{s_1s_2\cdots s_n}=\sum_{i_1+j_1=s_1}\sum_{i_2+j_2=s_2}\cdots\sum_{i_n+j_n=s_n}a_{i_1i_2\cdots i_n}b_{j_1j_2\cdots j_n}. \tag{4}$$

不难验证 $K[x_1,x_2,\cdots,x_n]$对于如上定义的加法与乘法成为一个环.它的零元素是零多项式.它有单位元素,即零次多项式 1.它是交换环,这个环称为数域 K 上的 n 元多项式环.

显然有

$$\deg(f+g)\le\max\{\deg f,\deg g\},$$

稍后我们将讨论 $\deg fg$ 与 $\deg f$,$\deg g$ 的关系.先来对 n 元多项式 $f(x_1,x_2,\cdots,x_n)$的各项规定一个排列顺序,从而给出首项的概念.

每一类单项式 $ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$ 对应一个 n 元有序非负整数组$(i_1,i_2,\cdots,i_n)$,这个对应是一一对应.为了给出各类单项式之间的一个排列顺序的方法,就只要对于 n 元有序非负整数组定义一个先后顺序就行了.

对于两个 n 元有序非负整数组$(i_1,i_2,\cdots,i_n)$与$(j_1,j_2,\cdots,j_n)$,如果 $i_1=$

$j_1, i_2=j_2,\cdots,i_{s-1}=j_{s-1}, i_s>j_s(1\leq s\leq n)$,则称$(i_1,i_2,\cdots,i_n)$先于$(j_1,j_2,\cdots,j_n)$,记作

$$(i_1,i_2,\cdots,i_n)>(j_1,j_2,\cdots,j_n).$$

由上述定义立即看出,对于任意两个 n 元有序非负整数组$(i_1,i_2,\cdots,i_n)$,$(j_1,j_2,\cdots,j_n)$,关系

$$(i_1,i_2,\cdots,i_n)>(j_1,j_2,\cdots,j_n),$$
$$(i_1,i_2,\cdots,i_n)=(j_1,j_2,\cdots,j_n),$$
$$(j_1,j_2,\cdots,j_n)>(i_1,i_2,\cdots,i_n)$$

中有一个且仅有一个成立.

关系"$>$"具有传递性,即如果

$$(i_1,i_2,\cdots,i_n)>(j_1,j_2,\cdots,j_n),$$

并且
$$(j_1,j_2,\cdots,j_n)>(k_1,k_2,\cdots,k_n),$$

则
$$(i_1,i_2,\cdots,i_n)>(k_1,k_2,\cdots,k_n).$$

这是因为 $i_1-k_1=(i_1-j_1)+(j_1-k_1)$,由此可得结论.例如,

$$(4,2,3,3)>(4,2,2,4),\quad (4,2,2,4)>(4,1,4,3),$$

由传递性得$(4,2,3,3)>(4,1,4,3)$.直接用定义也可看出$(4,2,3,3)>(4,1,4,3)$.

这样我们的确给出了 n 元有序非负整数组之间的一个顺序.相应地,n 元各类单项式之间也有了一个先后顺序.这种排列顺序的方法是模仿字典中单词的排列原则得出的,因而称之为字典排列法.

例如,多项式

$$2x_1^4x_2x_3+x_1x_2^5x_3+6x_1^3$$

按字典排列法写出来就是

$$2x_1^4x_2x_3+6x_1^3+x_1x_2^5x_3.$$

按字典排列法写出来的第一个系数不为零的单项式称为 n 元多项式的首项.例如,上面的多项式的首项是 $2x_1^4x_2x_3$.要注意,首项不一定具有最大的次数.例如,上面多项式的次数是 7,而首项的次数是 6.

定理1 在 $K[x_1,x_2,\cdots,x_n]$中两个非零多项式的乘积的首项等于它们的首项的乘积.从而两个非零多项式的乘积仍是非零多项式,即 $K[x_1,x_2,\cdots,x_n]$是无零因子环.

证明 设 $f(x_1,x_2,\cdots,x_n)$,$g(x_1,x_2,\cdots,x_n)$是 $K[x_1,x_2,\cdots,x_n]$中两个非零多项式.设 $f(x_1,x_2,\cdots,x_n)$的首项为 $ax_1^{p_1}x_2^{p_2}\cdots x_n^{p_n}$,$a\neq0$.设 $g(x_1,x_2,\cdots,x_n)$的首项是 $bx_1^{q_1}x_2^{q_2}\cdots x_n^{q_n}$,$b\neq0$.为了证明 fg 的首项是 $abx_1^{p_1+q_1}x_2^{p_2+q_2}\cdots x_n^{p_n+q_n}$,

只要证明

$$(p_1+q_1, p_2+q_2, \cdots, p_n+q_n).$$

先于 fg 中其他单项式的幂指数组就行了. fg 的其他单项式的幂指数组只有三种可能情形.

$$(p_1+j_1, p_2+j_2, \cdots, p_n+j_n),$$

或者
$$(i_1+q_1, i_2+q_2, \cdots, i_n+q_n),$$

或者
$$(i_1+j_1, i_2+j_2, \cdots, i_n+j_n),$$

其中
$$(p_1, p_2, \cdots, p_n) > (i_1, i_2, \cdots, i_n),$$
$$(q_1, q_2, \cdots, q_n) > (j_1, j_2, \cdots, j_n).$$

显然有

$$(p_1+q_1, p_2+q_2, \cdots, p_n+q_n) > (p_1+j_1, p_2+j_2, \cdots, p_n+j_n),$$
$$(p_1+q_1, p_2+q_2, \cdots, p_n+q_n) > (i_1+q_1, i_2+q_2, \cdots, i_n+q_n),$$
$$(i_1+q_1, i_2+q_2, \cdots, i_n+q_n) > (i_1+j_1, i_2+j_2, \cdots, i_n+j_n).$$

由传递性得

$$(p_1+q_1, p_2+q_2, \cdots, p_n+q_n) > (i_1+j_1, i_2+j_2, \cdots, i_n+j_n).$$

这就证明了 $abx_1^{p_1+q_1}x_2^{p_2+q_2}\cdots x_n^{p_n+q_n}$ 不可能与 fg 中其他的单项式相消,而且它先于 fg 其他的单项式,因此它是 fg 的首项. x

因为 $K[x_1, x_2, \cdots, x_n]$是无零因子环,所以消去律成立.

用数学归纳法可得出

推论 2　在 $K[x_1, x_2, \cdots, x_n]$中,如果 $f_i \neq 0$, $i=1,2,\cdots,m$,则 $f_1f_2\cdots f_m$ 的首项等于每个 f_i 的首项的乘积. x

我们要引进齐次多项式这一重要概念:

定义 2　数域 K 上的 n 元多项式 $g(x_1, x_2, \cdots, x_n)$称为 m 次齐次多项式,如果它的每个系数不为 0 的单项式都是 m 次的.

例如, $g(x_1, x_2, x_3) = 2x_1^4 + 3x_1^2x_2x_3 + x_1x_2x_3^2$ 是一个 4 次齐次多项式.

显然, $K[x_1, x_2, \cdots, x_n]$中两个齐次多项式的乘积仍是齐次多项式,它的次数等于这两个多项式的次数的和.

对于任何一个 n 元多项式 $f(x_1, x_2, \cdots, x_n)$,如果把 f 中所有次数相同的单项式并在一起,则 f 可以唯一地表示成

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=0}^{m} f_i(x_1, x_2, \cdots, x_n), \quad m = \deg f, \tag{5}$$

其中 $f_i(x_1, x_2, \cdots, x_n)$是 i 次齐次多项式,它称为 $f(x_1, x_2, \cdots, x_n)$的 i 次齐次成分.

利用(5)式我们可以证明下述定理.

定理3　设

$$f(x_1,x_2,\cdots,x_n),\quad g(x_1,x_2,\cdots,x_n)\in K[x_1,x_2,\cdots,x_n],$$

则
$$\deg fg=\deg f+\deg g. \tag{6}$$

证明　若 f,g 中有一个是零多项式,则(6)式成立.

现在设

$$f\neq 0,\quad g\neq 0,\quad \deg f=m,\quad \deg g=s,$$

则
$$f=f_0+f_1+\cdots+f_m,\quad g=g_0+g_1+\cdots+g_s,$$

其中 f_i 是 f 的 i 次齐次成分,g_j 是 g 的 j 次齐次成分,$f_m\neq 0$, $g_s\neq 0$. 我们有

$$\begin{aligned}fg=&f_0g_0+\cdots+f_0g_s+f_1g_0+\cdots+f_1g_s\\&+\cdots+f_mg_0+\cdots+f_mg_s,\end{aligned} \tag{7}$$

其中 f_ig_j 是 $i+j$ 次齐次多项式($i=0,1,\cdots,m$; $j=0,1,\cdots,s$). 因为 $f_m\neq 0$, $g_s\neq 0$,所以 $f_mg_s\neq 0$. 于是 f_mg_s 是 $m+s$ 次齐次多项式. 所以

$$\deg fg=m+s=\deg f+\deg g.$$
x

*当 $n>1$ 时, $K[x_1,x_2,\cdots,x_n]$中没有带余除法,但是唯一因式分解定理仍然成立.我们将在抽象代数课程中讨论这方面的问题.

在本章§1我们指出,数域 K 上的一元多项式环 $K[x]$有通用性质.现在我们要指出,数域 K 上的 n 元多项式环 $K[x_1,x_2,\cdots,x_n]$也具有通用性质,即

定理4　设 K 是一个数域,设 R 是一个有单位元的交换环,并且 R 可以看成是 K 的一个扩环(即 R 有一个子环 R_1 与 K 同构,把 R_1 与 K 的对应元素等同,并且 R 的单位元是 R_1 的单位元).设 $t_1,\cdots,t_n$ 是 R 的元素,令

$$\begin{aligned}\sigma_{t_1,\cdots,t_n}:K[x_1,x_2,\cdots,x_n]&\qquad R\\ f(x_1,x_2,\cdots,x_n)&\longmapsto f(t_1,t_2,\cdots,t_n),\end{aligned} \tag{8}$$

其中设

$$f(x_1,x_2,\cdots,x_n)=\sum_{i_1,i_2,\cdots,i_n}a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n},$$

而

$$f(t_1,t_2,\cdots,t_n)\overset{\mathrm{def}}{}\sum_{i_1,i_2,\cdots,i_n}a_{i_1i_2\cdots i_n}t_1^{i_1}t_2^{i_2}\cdots t_n^{i_n},$$

则 $\sigma_{t_1,\cdots,t_n}$是 $K[x_1,x_2,\cdots,x_n]$到 R 的一个映射,它使得

$$\sigma_{t_1,\cdots,t_n}(x_i)=t_i,\quad i=1,\cdots,n.$$

并且它保持加法与乘法运算,即如果

$$f(x_1,\cdots,x_n)+g(x_1,\cdots,x_n)=h(x_1,\cdots,x_n),$$
$$f(x_1,\cdots,x_n)g(x_1,\cdots,x_n)=p(x_1,\cdots,x_n),$$

那么 $$f(t_1,\cdots,t_n)+g(t_1,\cdots,t_n)=h(t_1,\cdots,t_n),$$
$$f(t_1,\cdots,t_n)g(t_1,\cdots,t_n)=p(t_1,\cdots,t_n).$$
由(8)式定义的映射 $\sigma_{t_1,\cdots,t_n}$ 称为 $x_1,\cdots,x_n$ 用 $t_1,\cdots,t_n$ 代入.

证明 与本章§1定理3的证明类似. x

$K[x_1,x_2,\cdots,x_n]$中所有零次多项式添上零多项式组成的子集是 $K[x_1,x_2,\cdots,x_n]$的一个子环,它与 K 是环同构的,因此 $K[x_1,x_2,\cdots,x_n]$可以看成是 K 的一个扩环.从而 $x_1,x_2,\cdots,x_n$ 可以用 $K[x_1,x_2,\cdots,x_n]$中任意 n 个元素代入,这种代入是保持加法与乘法运算的.

特别重要的一种情形是:$x_1,x_2,\cdots,x_n$ 用 K 中任意 n 个元素 $c_1,c_2,\cdots,c_n$ 代入,由此我们可引进多元多项式函数的概念:

设 $f(x_1,x_2,\cdots,x_n)$是数域 K 上的一个 n 元多项式,对于 K 中任意 n 个元素 $c_1,c_2,\cdots,c_n$,将 $x_1,x_2,\cdots,x_n$ 用 $c_1,c_2,\cdots,c_n$ 代入,得 $f(c_1,c_2,\cdots,c_n)\in K$. 于是 n 元多项式 $f(x_1,x_2,\cdots,x_n)$确定了集合 K^n 到数域 K 的一个映射(即 K 上的 n 元函数),仍用 f 表示,即

$$f:\ K^n \qquad\qquad K$$
$$(c_1,c_2,\cdots,c_n)\longmapsto f(c_1,c_2,\cdots,c_n). \tag{9}$$

这种由数域 K 上的 n 元多项式确定的 K 上的 n 元函数称为数域 K 上的 n 元多项式函数.

设 $f(x_1,x_2,\cdots,x_n),g(x_1,x_2,\cdots,x_n)$是数域 K 上的两个 n 元多项式.显然,如果 $f(x_1,x_2,\cdots,x_n)$与 $g(x_1,x_2,\cdots,x_n)$相等,则它们确定的 n 元多项式函数 f 与 g 也相等,反之如何?为此我们先证一个引理:

引理1 设 $h(x_1,\cdots,x_n)$是数域 K 上的一个 n 元多项式,如果 $h(x_1,\cdots,x_n)\neq 0$,则 h 不是零函数.

证明 对不定元的数目 n 作归纳法. $n=1$ 时,由于数域 K 上非零的一元多项式 $h(x)$给出的函数不是零函数,因此存在 $c\in K$,使 $h(c)\neq 0$.

假设命题对 $K[x_1,\cdots,x_{n-1}]$中的多项式成立,现在看 $K[x_1,\cdots,x_{n-1},x_n]$中的多项式 $h(x_1,\cdots,x_n)$.把 $h(x_1,\cdots,x_n)$写成

$$h(x_1,\cdots,x_n)=u_0(x_1,\cdots,x_{n-1})+u_1(x_1,\cdots,x_{n-1})x_n$$
$$+\cdots+u_s(x_1,\cdots,x_{n-1})x_n^s,$$

其中 $u_i(x_1,\cdots,x_{n-1})\in K[x_1,\cdots,x_{n-1}]$, $i=0,1,\cdots,s$,且 $u_s(x_1,\cdots,x_{n-1})\neq 0$.据归纳假设,$u_s$ 不是零函数,因此存在 $c_1,\cdots,c_{n-1}\in K$,使得 $u_s(c_1,\cdots,c_{n-1})\neq 0$.于是 $K[x_n]$中的多项式

$$h(c_1,\cdots,c_{n-1},x_n)=u_0(c_1,\cdots,c_{n-1})+u_1(c_1,\cdots,c_{n-1})x_n$$

$$+\cdots+u_s(c_1,\cdots,c_{n-1})x_n^s$$

是非零多项式,因此存在 $c_n\in K$,使得

$$h(c_1,\cdots,c_{n-1},c_n)=u_0(c_1,\cdots,c_{n-1})+u_1(c_1,\cdots,c_{n-1})c_n$$
$$+\cdots+u_s(c_1,\cdots,c_{n-1})c_n^s\neq 0.$$

x

注意:引理1证明中的关键是 K 有无穷多个元素.

从引理1立即得到

定理5　设 $f(x_1,\cdots,x_n),g(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$,如果多项式 $f(x_1,\cdots,x_n)$ 与 $g(x_1,\cdots,x_n)$ 不相等,则由它们确定的 n 元多项式函数 f 与 g 也不相等.

证明　考虑多项式

$$h(x_1,\cdots,x_n)=f(x_1,\cdots,x_n)-g(x_1,\cdots,x_n),$$

如果多项式 f 与 g 不相等,则 $h(x_1,\cdots,x_n)\neq 0$.据引理1,h 不是零函数,于是存在 $c_1,\cdots,c_n\in K$,使得 $h(c_1,\cdots,c_n)\neq 0$. $x_1,\cdots,x_n$ 用 $c_1,\cdots,c_n$ 代入,由上述式子可推出

$$f(c_1,\cdots,c_n)\neq g(c_1,\cdots,c_n),$$

所以映射 f 与 g 不相等.　x

我们把数域 K 上所有 n 元多项式函数组成的集合记作 K_{npol},在这个集合中规定加法与乘法如下:" $(c_1,\cdots,c_n)\in K^n$,

$$(f+g)(c_1,\cdots,c_n)\overset{\text{def}}{=}f(c_1,\cdots,c_n)+g(c_1,\cdots,c_n)\tag{10}$$

$$(fg)(c_1,\cdots,c_n)\overset{\text{def}}{=}f(c_1,\cdots,c_n)g(c_1,\cdots,c_n)\tag{11}$$

容易验证 K_{npol} 是一个环,称它为数域 K 上的 n 元多项式函数环.类似于§6的定理7的证法可以证明:数域 K 上的 n 元多项式环 $K[x_1,\cdots,x_n]$ 与 K 上的 n 元多项式函数环 K_{npol} 是同构的.因此我们可以把数域 K 上的 n 元多项式与 K 上的 n 元多项式函数等同看待.

设 $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$,对于 $c_1,\cdots,c_n\in K$,如果 $f(c_1,\cdots,c_n)=0$,则称 $(c_1,\cdots,c_n)$ 是 $f(x_1,\cdots,x_n)$ 的一个零点.当 K 取实数域,若 $n=2$,则 $f(x,y)$ 的零点组成的集合就是平面上的一条代数曲线;若 $n=3$,则 $f(x,y,z)$ 的零点组成的集合是空间中的一个代数曲面.研究数域 K 上一组 n 元多项式的公共零点组成的集合是代数几何的一个基本内容.

利用不定元 $x_1,\cdots,x_n$ 用 $K[x_1,\cdots,x_n]$ 的 n 个元素代入是保持运算的,我们可以得出齐次多项式的一个特征性质:

定理6　设 $f(x_1,\cdots,x_n)\in K[x_1,\cdots,x_n]$ 且 $f\neq 0$, m 是一个非负整数.则 $f(x_1,\cdots,x_n)$ 是 m 次齐次多项式的充分必要条件为对一切 $t\in K$,有

$$f(tx_1,tx_2,\cdots,tx_n)=t^m f(x_1,x_2,\cdots,x_n). \tag{12}$$

*证明 必要性.设 $f(x_1,\cdots,x_n)$是 m 次齐次多项式,即

$$f(x_1,\cdots,x_n)=\sum_{i_1,i_2,\cdots,i_n} a_{i_1i_2\cdots i_n}x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n},\quad i_1+i_2+\cdots+i_n=m,$$

任取 $t\in K$,不定元 $x_1,x_2,\cdots,x_n$ 用 $tx_1,tx_2,\cdots,tx_n$ 代入,从上式得

$$\begin{aligned}f(tx_1,\cdots,tx_n)&=\sum_{i_1,i_2,\cdots,i_n} a_{i_1i_2\cdots i_n}(tx_1)^{i_1}\cdots(tx_n)^{i_n}\\&=t^m f(x_1,\cdots,x_n).\end{aligned}$$

充分性.设对一切 $t\in K$,有(12)式成立.将 $f(x_1,\cdots,x_n)$写成

$$f=f_0+f_1+\cdots+f_s, \tag{13}$$

其中 f_i 是 f 的 i 次齐次成分,$i=0,1,\cdots,s$.任取 $t\in K$,$x_1,x_2,\cdots,x_n$ 用 $tx_1,tx_2,\cdots,tx_n$ 代入,从(13)式得出

$$\begin{aligned}f(tx_1,\cdots,tx_n)&=f_0(tx_1,\cdots,tx_n)+f_1(tx_1,\cdots,tx_n)\\&\quad+\cdots+f_s(tx_1,\cdots,tx_n).\end{aligned}$$

据刚证的必要性以及充分性的假设得

$$\begin{aligned}t^m f(x_1,\cdots,x_n)&=f_0(x_1,\cdots,x_n)+tf_1(x_1,\cdots,x_n)\\&\quad+\cdots+t^s f_s(x_1,\cdots,x_n).\end{aligned} \tag{14}$$

将(13)式代入(14)式左端,并且根据 $K[x_1,\cdots,x_n]$中两个多项式相等的规定,我们得到

$$t^m f_i(x_1,\cdots,x_n)=t^i f_i(x_1,\cdots,x_n),\quad i=0,1,\cdots,s. \tag{15}$$

任取 $i\in\{0,1,\cdots,s\}$且 $i\neq m$.如果 $f_i\neq 0$,则从(15)式两边消去 f_i 得 $t^m=t^i$,$\forall t\in K$.据§6的定理6得,$x^m=x^i$,这与 $i\neq m$ 矛盾.因此 $f_i=0$,当 $i\neq m$.所以 $f=f_m$.已知 $f\neq 0$,因此 f 是 m 次齐次多项式. □

我们指出,设 R 是(或可看成是)K 的一个交换扩环,任取 R 中 n 个元素 $t_1,t_2,\cdots,t_n$,对于 $f(x_1,x_2,\cdots,x_n)\in K[x_1,x_2,\cdots,x_n]$,把 $x_1,x_2,\cdots,x_n$ 用 $t_1,t_2,\cdots,t_n$ 代入,得到的 $f(t_1,t_2,\cdots,t_n)$称为 $t_1,t_2,\cdots,t_n$ 在 K 上的一个多项式.尽管从形式上看,$t_1,t_2,\cdots,t_n$ 的多项式 $f(t_1,t_2,\cdots,t_n)$与 n 元多项式 $f(x_1,x_2,\cdots,x_n)$类似,但它们有本质不同:$t_1,t_2,\cdots,t_n$ 的一个多项式,其表法可能不唯一,而一个 n 元多项式,其表法唯一.

习 题 7.9

1. 将下列4元多项式按字典排列法排列各单项式的顺序:

(1) $f(x_1, x_2, x_3, x_4) = x_3^4 x_4 - x_1^3 x_2 + 5 x_2 x_3 x_4 + 2 x_2^4 x_3 x_4$;

(2) $f(x_1, x_2, x_3, x_4) = x_1^3 + x_3^2 + 3 x_1 x_2^2 x_4 - 5 x_1^2 x_3 x_4^2 - 2 x_2^3 x_3$.

2. 把3元齐次多项式 $f(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3 - 3 x_1 x_2 x_3$ 写成两个3元齐次多项式的乘积.

3. 设 $f, g \in K[x_1, \cdots, x_n]$,且 $g \neq 0$.证明:如果对于使 $g(c_1, \cdots, c_n) \neq 0$ 的任意一组元素 $c_1, \cdots, c_n \in K$,都有 $f(c_1, \cdots, c_n) = 0$,则 $f(x_1, \cdots, x_n) = 0$.

*4. 设 $f, g \in K[x_1, \cdots, x_n]$,如果存在 $h \in K[x_1, \cdots, x_n]$,使得 $f = gh$,则称 g 整除 f,记作 $g | f$.此时称 g 是 f 的一个因式.如果 $f | g$,并且 $g | f$,则称 f 与 g 是相伴的,记作 $f \sim g$.

(1) 证明: $f \sim g$　　　　f 与 g 相差一个非零数因子;

(2) $f \in K[x_1, \cdots, x_n]$称为不可约的,如果 f 的因式只有零次多项式以及 f 的相伴元.否则称为可约的.证明:在 $K[x, y]$ 里,多项式 $x^2 - y$ 是不可约的.

§10　对称多项式

这一节我们讨论一类重要的多元多项式,即对称多项式.在给出对称多项式的概念之前,先介绍置换的概念.

定义1　一个有限集合 Ω 到自身的一个双射叫做 Ω 的一个置换.

设 Ω 是具有 n 个元素的有限集.因为我们讨论的问题与这些元素本身无关,所以我们不妨假定 $\Omega = \{1, 2, \cdots, n\}$.设 σ 是 Ω 的一个置换(也称 σ 为 n 元置换),设 i 在映射 σ 下的象是 a_i,即

$$\sigma(i) = a_i, \qquad i = 1, 2, \cdots, n.$$

我们可以把置换 σ 表示如下:

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}. \tag{1}$$

由于 σ 是 Ω 到自身的一个双射,所以 $a_1 a_2 \cdots a_n$ 是自然数 $1, 2, \cdots, n$ 的一个全排列,即 n 元排列.反之,如果给了一个 n 元排列 $a_1 a_2 \cdots a_n$,则 $\sigma: i \longmapsto a_i$ 是 Ω 的一个置换.因此 n 元置换组成的集合与 n 元排列组成的集合之间有一个一一对应.我们用 S_n 表示所有 n 元置换组成的集合,由刚才的一一对应知道,$|S_n| = n!$.

例如,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix} \tag{2}$$

是一个4元置换,$\sigma(1) = 3$, $\sigma(2) = 1$, $\sigma(3) = 4$, $\sigma(4) = 2$.

设 σ 是 Ω 的一个置换,因为 σ 是 Ω 到自身的一个双射,所以 σ 是可逆映射.

σ的逆映射σ^{-1}称为置换σ的逆置换.例如,(2)式给出的置换σ的逆置换σ^{-1}如下:

$$\sigma^{-1}(3)=1,\quad \sigma^{-1}(1)=2,\quad \sigma^{-1}(4)=3,\quad \sigma^{-1}(2)=4$$

于是σ^{-1}可以表示成

$$\sigma^{-1}=\begin{pmatrix}1&2&3&4\\2&4&1&3\end{pmatrix}. \tag{3}$$

请读者直接从(2)式和(3)式去验证$\sigma^{-1}\sigma=1_\Omega$,$\sigma\sigma^{-1}=1_\Omega$,其中$1_\Omega$表示$\Omega$到自身的恒等映射.

如果一个n元置换把i映成j,把j映成i,而其余元素不变,则称这个置换是一个对换,记作(ij).显然$(ij)^{-1}=(ij)$.

任意给定一个$\sigma\in S_n$,由于$K[x_1,\cdots,x_n]$可以看成是K的一个扩环,据n元多项式环的通用性质,$x_1,x_2,\cdots,x_n$用$x_{\sigma^{-1}(1)},x_{\sigma^{-1}(2)},\cdots,x_{\sigma^{-1}(n)}$代入是$K[x_1,x_2,\cdots,x_n]$到$K[x_1,x_2,\cdots,x_n]$的一个映射,并且它保持加法与乘法运算,这个映射记作σ.即

$$\begin{aligned}\sigma: K[x_1,x_2,\cdots,x_n] &\to K[x_1,x_2,\cdots,x_n]\\ f(x_1,x_2,\cdots,x_n)&\longmapsto f(x_{\sigma^{-1}(1)},x_{\sigma^{-1}(2)},\cdots,x_{\sigma^{-1}(n)}),\end{aligned}$$

也就是说

$$(\sigma f)(x_1,x_2,\cdots,x_n)=f(x_{\sigma^{-1}(1)},x_{\sigma^{-1}(2)},\cdots,x_{\sigma^{-1}(n)}). \tag{4}$$

我们把上述σ称为置换σ诱导的代入.

观察2元多项式

$$f(x_1,x_2)=3x_1^2+3x_2^2+4x_1x_2,$$

把x_1,x_2的下标对换得到多项式$3x_2^2+3x_1^2+4x_2x_1$,它与原来的多项式$f(x_1,x_2)$相等.像这样的多项式称为对称多项式.

定义2　设$f(x_1,x_2,\cdots,x_n)\in K[x_1,x_2,\cdots,x_n]$,如果对一切$\sigma\in S_n$都有$\sigma f=f$,则称$f(x_1,x_2,\cdots,x_n)$是数域$K$上的一个$n$元对称多项式.

从定义2得出,如果一个n元对称多项式$f(x_1,x_2,\cdots,x_n)$含有一项$ax_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$,那么它也含有项

$$ax_{\sigma^{-1}(1)}^{i_1}x_{\sigma^{-1}(2)}^{i_2}\cdots x_{\sigma^{-1}(n)}^{i_n},\qquad \forall\,\sigma\in S_n,$$

由于当σ跑遍一切n元置换时,$\sigma^{-1}(1)\sigma^{-1}(2)\cdots\sigma^{-1}(n)$便取遍$1,2,\cdots,n$的所有$n$元排列,因此$f(x_1,x_2,\cdots,x_n)$含有一切形如

$$ax_{j_1}^{i_1}x_{j_2}^{i_2}\cdots x_{j_n}^{i_n}$$

的项,其中$j_1j_2\cdots j_n$是$1,2,\cdots,n$的任一n元排列.

例如,若3元对称多项式$f(x_1,x_2,x_3)$含有一项$x_1^2x_2$,则$f(x_1,x_2,x_3)$应

当含有一切形如 $x_{j_1}^2 x_{j_2} x_{j_3}^0$ 的项,其中 $j_1 j_2 j_3$ 是 1,2,3 的任一 3 元排列.从而 $f(x_1, x_2, x_3)$还应当含有下列各项:

$$x_1^2 x_3, \quad x_2^2 x_1, \quad x_2^2 x_3, \quad x_3^2 x_1, \quad x_3^2 x_2.$$

于是 3 元对称多项式中,含有 $x_1^2 x_2$ 的项数最少的对称多项式为

$$x_1^2 x_2 + x_1^2 x_3 + x_2^2 x_1 + x_2^2 x_3 + x_3^2 x_1 + x_3^2 x_2.$$

又如,零多项式和零次多项式都是对称多项式.

现在我们来研究 $K[x_1, x_2, \cdots, x_n]$中所有对称多项式组成的集合 W 的结构.

命题 1　W 是 $K[x_1, x_2, \cdots, x_n]$的一个子环.

证明　显然 W 非空集.为了证 W 是子环,只要证 W 对于减法和乘法封闭.任取 $f, g \in W$,则对于任意 $\sigma \in S_n$,有 $\sigma f = f, \sigma g = g$.因为 σ 是保持加法与乘法运算的,所以有

$$\sigma(f - g) = \sigma f - \sigma g = f - g,$$
$$\sigma(fg) = (\sigma f)(\sigma g) = fg.$$

从而 $f - g, fg \in W$.因此 W 是 $K[x_1, x_2, \cdots, x_n]$的一个子环.　x

命题 1 表明, $K[x_1, x_2, \cdots, x_n]$中的对称多项式的和、差、积仍是对称多项式.

我们下一个任务是研究这个子环 W 的结构.

命题 2　设 $f_1, f_2, \cdots, f_n \in W$,则对于 $K[x_1, x_2, \cdots, x_n]$中任意一个多项式 $g(x_1, x_2, \cdots, x_n)$,都有

$$h(x_1, x_2, \cdots, x_n) \overset{\text{def}}{=} g(f_1, f_2, \cdots, f_n)$$

是对称多项式.

证明　设

$$g(x_1, x_2, \cdots, x_n) = \sum_{i_1, i_2, \cdots, i_n} a_{i_1 i_2 \cdots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n},$$

则

$$h(x_1, x_2, \cdots, x_n) \overset{\text{def}}{=} g(f_1, f_2, \cdots, f_n)$$
$$= \sum_{i_1, i_2, \cdots, i_n} a_{i_1 i_2 \cdots i_n} f_1^{i_1} f_2^{i_2} \cdots f_n^{i_n} \in W.$$　x

命题 2 表明,如果知道了 $K[x_1, x_2, \cdots, x_n]$中的 n 个对称多项式 $f_1, f_2, \cdots, f_n$,那么我们可以得到无穷多个对称多项式,自然会问:能不能找到 n 个最基本的对称多项式,使得 $K[x_1, x_2, \cdots, x_n]$中的每一个对称多项式都能通过它们用命题 2 的方法得到?注意数域 K 上首项系数为 1 的一元多项式 $f(x)$的 n 个复

根 $c_1, c_2, \cdots, c_n$ 与 $f(x)$ 的系数的 n 个关系式中,每一个表达式都是关于 $c_1, c_2, \cdots, c_n$ 对称的. 由此受到启发,考虑下列 n 个 n 元多项式:

$$
\begin{aligned}
&\sigma_1(x_1, x_2, \cdots, x_n) = \sum_{j=1}^{n} x_j, \\
&\sigma_2(x_1, x_2, \cdots, x_n) = \sum_{1 \le j_1 < j_2 \le n} x_{j_1} x_{j_2}, \\
&\cdots\cdots\cdots\cdots \\
&\sigma_k(x_1, x_2, \cdots, x_n) = \sum_{1 \le j_1 < j_2 < \cdots < j_k \le n} x_{j_1} x_{j_2} \cdots x_{j_k}, \\
&\cdots\cdots\cdots\cdots \\
&\sigma_n(x_1, x_2, \cdots, x_n) = x_1 x_2 \cdots x_n.
\end{aligned} \tag{5}
$$

这 n 个多项式都是对称多项式,理由如下:设一个 n 元对称多项式含有一项 $x_1 x_2 \cdots x_k$,则它应当含有一切形如

$$x_{j_1} x_{j_2} \cdots x_{j_k} x_{j_{k+1}}^0 \cdots x_{j_n}^0$$

的项,其中 $j_1 j_2 \cdots j_k j_{k+1} \cdots j_n$ 是 $1,2,\cdots,n$ 的任一 n 元排列.

因此 $\sigma_k(x_1, x_2, \cdots, x_n) = \sum_{1 \le j_1 < \cdots < j_k \le n} x_{j_1} x_{j_2} \cdots x_{j_k}$ 是对称多项式, $k = 1,2,\cdots,n$.

$\sigma_k(x_1, x_2, \cdots, x_n)$, $k = 1, 2, \cdots, n$,称为初等对称多项式. 它们正是我们要找的 n 个最基本的对称多项式. 即,我们有下述关于子环 W 的结构的重要定理:

定理 3(对称多项式基本定理) 对于 $K[x_1, x_2, \cdots, x_n]$ 中任意一个对称多项式 $f(x_1, x_2, \cdots, x_n)$,都存在 $K[x_1, x_2, \cdots, x_n]$ 中唯一的一个多项式 $g(x_1, x_2, \cdots, x_n)$,使得

$$f(x_1, x_2, \cdots, x_n) = g(\sigma_1, \sigma_2, \cdots, \sigma_n).$$

证明 存在性. 采取首项消去法. 设对称多项式 $f(x_1, x_2, \cdots, x_n)$ 的首项是

$$a x_1^{l_1} x_2^{l_2} \cdots x_n^{l_n}, \quad a \neq 0. \tag{6}$$

我们断言 $l_1 \ge l_2 \ge \cdots \ge l_n$. 理由如下:假如 $l_i < l_{i+1}$,由于 $f(x_1, x_2, \cdots, x_n)$ 是对称多项式,因此对换 $\tau = (i, i+1)$ 诱导的代入 τ 使得 $\tau f = f$. 于是项(6)在 τ 下的象

$$a x_1^{l_1} x_2^{l_2} \cdots x_i^{l_{i+1}} x_{i+1}^{l_i} \cdots x_n^{l_n} \tag{7}$$

也是 f 的一项. 而项(7)的幂指数组 $(l_1, \cdots, l_{i-1}, l_{i+1}, l_i, \cdots, l_n)$ 先于项(6)的幂指数组 $(l_1, \cdots, l_{i-1}, l_i, l_{i+1}, \cdots, l_n)$,这与项(6)是 $f(x_1, x_2, \cdots, x_n)$ 的首项矛盾. 所以 $l_1 \ge l_2 \ge \cdots \ge l_n$.

为了消去 $f(x_1, x_2, \cdots, x_n)$ 的首项,同时又要出现 $\sigma_1, \sigma_2, \cdots, \sigma_n$,我们作多

项式

$$\varphi_1(x_1, x_2, \cdots, x_n) = a\sigma_1^{l_1 - l_2}\sigma_2^{l_2 - l_3}\cdots\sigma_{n-1}^{l_{n-1} - l_n}\sigma_n^{l_n}, \tag{8}$$

因为 $K[x_1, x_2, \cdots, x_n]$ 中对称多项式的乘积还是对称多项式,所以 $\varphi_1(x_1, x_2, \cdots, x_n)$ 是对称多项式.又由于多项式的乘积的首项等于它们的首项的乘积,因此 $\varphi_1(x_1, x_2, \cdots, x_n)$ 的首项是

$$\begin{aligned} & ax_1^{l_1 - l_2}(x_1x_2)^{l_2 - l_3}\cdots(x_1x_2\cdots x_{n-1})^{l_{n-1} - l_n}(x_1x_2\cdots x_{n-1}x_n)^{l_n} \\ = & ax_1^{l_1}x_2^{l_2}\cdots x_{n-1}^{l_{n-1}}x_n^{l_n}, \end{aligned}$$

它等于 $f(x_1, x_2, \cdots, x_n)$ 的首项.令

$$\begin{aligned} & f_1(x_1, x_2, \cdots, x_n) \\ = & f(x_1, x_2, \cdots, x_n) - \varphi_1(x_1, x_2, \cdots, x_n), \end{aligned} \tag{9}$$

则 f_1 的首项"小于"f 的首项(即 f 的首项的幂指数组先于 f_1 的首项的幂指数组),并且由于对称多项式的差仍是对称多项式,所以 $f_1(x_1, x_2, \cdots, x_n)$ 是 $K[x_1, x_2, \cdots, x_n]$ 中的对称多项式.

对 $f_1(x_1, x_2, \cdots, x_n)$ 重复上述作法,我们又得到 $K[x_1, x_2, \cdots, x_n]$ 中的一个对称多项式

$$f_2(x_1, x_2, \cdots, x_n) = f_1(x_1, x_2, \cdots, x_n) - \varphi_2(x_1, x_2, \cdots, x_n),$$

其中 $\varphi_2(x_1, x_2, \cdots, x_n)$ 是 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的适当方幂的乘积,并且其系数等于 f_1 的首项系数(类似于(8)),而 f_2 的首项"小于"f_1 的首项.

如此继续下去,我们得到 $K[x_1, x_2, \cdots, x_n]$ 中一系列的对称多项式

$$f, \quad f_1 = f - \varphi_1, \quad f_2 = f_1 - \varphi_2, \quad \cdots \tag{10}$$

它们的首项的幂指数组一个比一个"小",其中 φ_i 是 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 适当方幂的乘积,并且其系数等于 f_{i-1} 的首项系数.设 f_i 的首项的幂指数组为 $(p_1, p_2, \cdots, p_n)$,则

$$(l_1, l_2, \cdots, l_n) > (p_1, p_2, \cdots, p_n),$$

于是 $l_1 \geq p_1$.又因为 f_i 是对称多项式,所以 $p_1 \geq p_2 \geq \cdots \geq p_n$.因此 $l_1 \geq p_1 \geq p_2 \geq \cdots \geq p_n$.满足这个条件的非负整数组 $(p_1, p_2, \cdots, p_n)$ 只有有限多个,因此序列(10)中只能有有限多个 f_i 不为零.即有正整数 s,使得 $f_s = 0$.于是有

$$f_1 = f - \varphi_1, \quad f_2 = f_1 - \varphi_2, \quad \cdots,$$

$$f_{s-1} = f_{s-2} - \varphi_{s-1}, \quad f_s = f_{s-1} - \varphi_s = 0,$$

从而得到

$$f = \varphi_1 + \varphi_2 + \cdots + \varphi_s. \tag{11}$$

设 $\varphi_i(x_1, x_2, \cdots, x_n) = a_i\sigma_1^{t_{i1}}\sigma_2^{t_{i2}}\cdots\sigma_n^{t_{in}}$,令

$$g(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{s} a_i x_1^{t_{i1}}x_2^{t_{i2}}\cdots x_n^{t_{in}} \in K[x_1, x_2, \cdots, x_n], \tag{12}$$

则 $$g(\sigma_1,\sigma_2,\cdots,\sigma_n)=\sum_{i=1}^{s}a_i\sigma_1^{t_{i1}}\sigma_2^{t_{i2}}\cdots\sigma_n^{t_{in}}=\sum_{i=1}^{s}\phi_i(x_1,x_2,\cdots,x_n)$$
$$=f(x_1,x_2,\cdots,x_n).$$

存在性证毕.

*唯一性.如果 $K[x_1,x_2,\cdots,x_n]$中有两个不同的多项式 $g_1(x_1,x_2,\cdots,x_n)$, $g_2(x_1,x_2,\cdots,x_n)$,使得

$$f(x_1,x_2,\cdots,x_n)=g_1(\sigma_1,\sigma_2,\cdots,\sigma_n),$$
$$f(x_1,x_2,\cdots,x_n)=g_2(\sigma_1,\sigma_2,\cdots,\sigma_n),$$

则 $$g_1(\sigma_1,\sigma_2,\cdots,\sigma_n)-g_2(\sigma_1,\sigma_2,\cdots,\sigma_n)=0.$$

令 $$g(x_1,x_2,\cdots,x_n)$$
$$=g_1(x_1,x_2,\cdots,x_n)-g_2(x_1,x_2,\cdots,x_n), \quad (13)$$

则 $$g(\sigma_1,\sigma_2,\cdots,\sigma_n)$$
$$=g_1(\sigma_1,\sigma_2,\cdots,\sigma_n)-g_2(\sigma_1,\sigma_2,\cdots,\sigma_n)=0. \quad (14)$$

由假设以及(13)式得

$$g(x_1,x_2,\cdots,x_n)\neq 0,$$

于是据§9的引理1得,存在 $b_1,b_2,\cdots,b_n\in K$,使得

$$g(b_1,b_2,\cdots,b_n)\neq 0. \quad (15)$$

令 $$\phi(x)=x^n-b_1x^{n-1}+\cdots+(-1)^kb_kx^{n-k}+\cdots+(-1)^nb_n,$$

设 $\phi(x)$的 n 个复根是 $\alpha_1,\alpha_2,\cdots,\alpha_n$,则从 Vieta 公式推出

$$b_1=\sigma_1(\alpha_1,\alpha_2,\cdots,\alpha_n),\cdots,b_k=\sigma_k(\alpha_1,\alpha_2,\cdots,\alpha_n),\cdots,b_n=\sigma_n(\alpha_1,\alpha_2,\cdots,\alpha_n), \quad (16)$$

$x_1,x_2,\cdots,x_n$ 用 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 代入,从(14)式得

$$g(\sigma_1(\alpha_1,\cdots,\alpha_n),\cdots,\sigma_n(\alpha_1,\cdots,\alpha_n))=0,$$

即 $$g(b_1,b_2,\cdots,b_n)=0.$$

矛盾.唯一性得证. x

对称多项式基本定理完全解决了 $K[x_1,x_2,\cdots,x_n]$中所有对称多项式组成的子环 W 的结构问题.定理中存在性的证明是构造性的,可以实际地利用它去求多项式 $g(x_1,x_2,\cdots,x_n)$,使得

$$f(x_1,x_2,\cdots,x_n)=g(\sigma_1,\sigma_2,\cdots,\sigma_n).$$

例 1 在 $K[x_1,x_2,x_3]$中,用初等对称多项式表出对称多项式

$$f(x_1,x_2,x_3)=x_1^2x_2^2+x_1^2x_3^2+x_2^2x_3^2.$$

解 $f(x_1,x_2,x_3)$的首项是 $x_1^2x_2^2$,它的幂指数组为(2,2,0).作多项式

$$\phi_1(x_1,x_2,x_3)=\sigma_1^{2-2}\sigma_2^{2-0}\sigma_3^0=\sigma_2^2,$$

令

$$
\begin{aligned}
& f_1(x_1, x_2, x_3) \\
&= f(x_1, x_2, x_3) - \varphi_1(x_1, x_2, x_3) \\
&= x_1^2 x_2^2 + x_1^2 x_3^2 + x_2^2 x_3^2 - (x_1 x_2 + x_1 x_3 + x_2 x_3)^2 \\
&= -2(x_1^2 x_2 x_3 + x_1 x_2^2 x_3 + x_1 x_2 x_3^2) \\
&= -2\sigma_1 \sigma_3,
\end{aligned}
$$

$f_1(x_1, x_2, x_3)$的首项是$-2x_1^2 x_2 x_3$,它的幂指数组为(2,1,1).令

$$\varphi_2(x_1, x_2, x_3) = -2\sigma_1^{2-1}\sigma_2^{1-1}\sigma_3^{1} = -2\sigma_1\sigma_3,$$

$$f_2(x_1, x_2, x_3) = f_1(x_1, x_2, x_3) - \varphi_2(x_1, x_2, x_3) = 0,$$

于是

$$f(x_1, x_2, x_3) = \sigma_2^2 - 2\sigma_1\sigma_3.$$

对于较复杂的 n 元对称多项式 $f(x_1, x_2, \cdots, x_n)$,求一个多项式 $g(x_1, x_2, \cdots, x_n)$,使得 $f(x_1, x_2, \cdots, x_n) = g(\sigma_1, \sigma_2, \cdots, \sigma_n)$,采用待定系数法更为简便.我们举一个例子来说明这种方法.

例 2　在 $K[x_1, x_2, \cdots, x_n]$中,用初等对称多项式表出对称多项式

$$f(x_1, x_2, \cdots, x_n) = \sum x_1^2 x_2^2,$$

这里$\sum x_1^2 x_2^2$ 表示 $x_1^2 x_2^2$ 经过 S_n 中任一置换诱导的代入所得到的项的和.

解　基本定理证明中已指出,序列(10)中多项式 f_i 的首项的幂指数组$(p_1, p_2, \cdots, p_n)$应当满足

$$l_1 \geq p_1 \geq p_2 \geq \cdots \geq p_n,$$

其中$(l_1, l_2, \cdots, l_n)$是 f 的首项的幂指数组.本题 f 的首项 $x_1^2 x_2^2$ 的幂指数组为(2,2,0,…,0).所以$(p_1, p_2, \cdots, p_n)$应满足

$$2 \geq p_1 \geq p_2 \geq \cdots \geq p_n.$$

因为初等对称多项式 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 都是齐次多项式,而齐次多项式的乘积仍是齐次多项式,所以基本定理证明中的 φ_i 都是齐次多项式.由于 φ_1 的首项等于 f 的首项 $x_1^2 x_2^2$,所以 φ_1 是 4 次齐次多项式.本题的 f 显然是 4 次齐次多项式.于是 $f_1 = f - \varphi_1$ 也是 4 次齐次多项式.同理 $f_2 = f_1 - \varphi_2$ 也是 4 次齐次多项式.依次下去,对于本题,序列(10)中非零的 f_i 都是 4 次齐次多项式.因此应当有

$$p_1 + p_2 + \cdots + p_n = 4.$$

从上两段可知,可能的幂指数组只有三个:

$$(2,2,0,\cdots,0),\quad (2,1,1,0,\cdots,0),\quad (1,1,1,1,0,\cdots,0),$$

它们分别是 f, f_1, f_2 的首项的幂指数组.由 φ 的构造方法得

$$\varphi_1 = \sigma_1^{2-2}\sigma_2^{2-0}\sigma_3^{0-0}\cdots\sigma_{n-1}^{0-0}\sigma_n^{0} = \sigma_2^2,$$

$$\varphi_2 = a\sigma_1^{2-1}\sigma_2^{1-1}\sigma_3^{1-0}\sigma_4^{0-0}\cdots\sigma_{n-1}^{0-0}\sigma_n^{0} = a\sigma_1\sigma_3,$$

$$\varphi_3 = b\sigma_1^{1-1}\sigma_2^{1-1}\sigma_3^{1-1}\sigma_4^{1-0}\sigma_5^{0-0}\cdots\sigma_{n-1}^{0-0}\sigma_n^{0} = b\sigma_4,$$

其中 a,b 是待定的系数.据基本定理的证明,得

$$f(x_1,x_2,\cdots,x_n)=\sigma_2^2+a\sigma_1\sigma_3+b\sigma_4. \tag{17}$$

为了决定 a,b,先让 $x_1,x_2,x_3,x_4,\cdots,x_n$ 用 $1,1,1,0,\cdots,0$ 代入,f 的值为 3,$\sigma_1,\sigma_2,\sigma_3,\sigma_4$ 的值依次是 3,3,1,0,于是从(17)式得到

$$3=3^2+a\cdot 3\cdot 1+b\cdot 0,$$

解得 $a=-2$.于是(17)式变成

$$f(x_1,x_2,\cdots,x_n)=\sigma_2^2-2\sigma_1\sigma_3+b\sigma_4. \tag{18}$$

$x_1,x_2,x_3,x_4,\cdots,x_n$ 用 $1,1,1,1,0,\cdots,0$ 代入,从(18)式得

$$6=6^2-2\cdot 4\cdot 4+b\cdot 1,$$

解得 $b=2$.于是

$$f(x_1,x_2,\cdots,x_n)=\sigma_2^2-2\sigma_1\sigma_3+2\sigma_4.$$

例 2 中的对称多项式 $f(x_1,x_2,\cdots,x_n)$是齐次多项式.如果所给的对称多项式 f 不是齐次的,那么把 f 表示成它的齐次成分的和.对每一个齐次成分(它仍然是对称多项式)采用例 2 的做法,最后把所得的结果相加即可

对称多项式基本定理的一个重要应用是,研究数域 K 上一元多项式在复数域中有无重根.

设数域 K 上首项系数为 1 的多项式

$$f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

在复数域中的 n 个根为 $c_1,c_2,\cdots,c_n$.容易看出

$$f(x)\text{在复数域中有重根} \Leftrightarrow D(c_1,c_2,\cdots,c_n)\overset{\text{def}}{=}\prod_{1\leq j<i\leq n}(c_i-c_j)^2=0.$$

由于 $f(x)$的 n 个复根 $c_1,c_2,\cdots,c_n$ 是不知道的,因此想用 $f(x)$的系数来表示 $D(c_1,c_2,\cdots,c_n)$.据 Vieta 公式得

$$\begin{aligned}-a_{n-1}&=c_1+c_2+\cdots+c_n=\sigma_1(c_1,c_2,\cdots,c_n),\\ a_{n-2}&=\sum_{1\leq j_1<j_2\leq n}c_{j_1}c_{j_2}=\sigma_2(c_1,c_2,\cdots,c_n),\\ &\cdots\cdots\cdots\cdots\\ (-1)^na_0&=c_1c_2\cdots c_n=\sigma_n(c_1,c_2,\cdots,c_n).\end{aligned} \tag{19}$$

由此受到启发,如果 $D(c_1,c_2,\cdots,c_n)$能够用 $\sigma_1(c_1,c_2,\cdots,c_n),\sigma_2(c_1,c_2,\cdots,c_n),\cdots,\sigma_n(c_1,c_2,\cdots,c_n)$表示出来,那么它就能够用 $f(x)$的系数 a_{n-1},$a_{n-2},\cdots,a_0$ 表示出来.注意到 $D(c_1,c_2,\cdots,c_n)=\prod_{1\leq j<i\leq n}(c_i-c_j)^2$ 是关于 c_1,$c_2,\cdots,c_n$ 对称的表达式,因此自然会想到运用对称多项式基本定理.

考虑数域 K 上 n 元对称多项式

$$D(x_1, x_2, \cdots, x_n) = \prod_{1 \le j < i \le n} (x_i - x_j)^2. \tag{20}$$

根据对称多项式基本定理,存在 $K[x_1, x_2, \cdots, x_n]$中唯一的一个多项式 $g(x_1, x_2, \cdots, x_n)$,使得

$$D(x_1, x_2, \cdots, x_n) = g(\sigma_1, \sigma_2, \cdots, \sigma_n). \tag{21}$$

不定元 $x_1, x_2, \cdots, x_n$ 分别用 $c_1, c_2, \cdots, c_n$ 代入,从(19)、(21)式得

$$D(c_1, c_2, \cdots, c_n) = g(-a_{n-1}, a_{n-2}, \cdots, (-1)^n a_0). \tag{22}$$

于是得出

命题4　数域 K 上首项系数为1的 n 次多项式

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

在复数域中有重根的充分必要条件为

$$g(-a_{n-1}, a_{n-2}, \cdots, (-1)^n a_0) = 0. \tag{23}$$

我们把 $f(x)$的系数 $a_{n-1}, a_{n-2}, \cdots, a_0$ 的多项式　　x

$$g(-a_{n-1}, a_{n-2}, \cdots, (-1)^n a_0)$$

称为 $f(x)$的判别式,记作 $D(f)$.利用它可以判断 $f(x)$在复数域中有无重根.

如何求出 $f(x)$的判别式 $D(f)$呢?由(22)式得

$$\begin{aligned} D(f) &= g(-a_{n-1}, a_{n-2}, \cdots, (-1)^n a_0) = D(c_1, c_2, \cdots, c_n) \\ &= \prod_{1 \le j < i \le n} (c_i - c_j)^2. \end{aligned} \tag{24}$$

表达式 $\prod_{1 \le j < i \le n} (c_i - c_j)^2$ 使人联想到范德蒙行列式,令

$$B = \begin{matrix} 1 & 1 & \cdots & 1 \\ c_1 & c_2 & \cdots & c_n \\ \cdots & \cdots & & \cdots \\ c_1^{n-1} & c_2^{n-1} & \cdots & c_n^{n-1} \end{matrix}, \tag{25}$$

则$|B| = \prod_{1 \le j < i \le n} (c_i - c_j)$.又由于$|B'| = |B|$,因此从(24)式得

$$D(f) = \prod_{1 \le j < i \le n} (c_i - c_j)^2 = |B||B'| = |BB'|. \tag{26}$$

于是

$$D(f) = \left| \begin{matrix} 1 & 1 & \cdots & 1 \\ c_1 & c_2 & \cdots & c_n \\ \cdots & \cdots & & \cdots \\ c_1^{n-1} & c_2^{n-1} & \cdots & c_n^{n-1} \end{matrix} \quad \begin{matrix} 1 & c_1 & \cdots & c_1^{n-1} \\ 1 & c_2 & \cdots & c_2^{n-1} \\ \cdots & \cdots & & \cdots \\ 1 & c_n & \cdots & c_n^{n-1} \end{matrix} \right|$$

$$= \begin{vmatrix} n & \sum_{i=1}^{n} c_i & \cdots & \sum_{i=1}^{n} c_i^{n-1} \\ \sum_{i=1}^{n} c_i & \sum_{i=1}^{n} c_i^2 & \cdots & \sum_{i=1}^{n} c_i^{n} \\ \cdots & \cdots & & \cdots \\ \sum_{i=1}^{n} c_i^{n-1} & \sum_{i=1}^{n} c_i^{n} & \cdots & \sum_{i=1}^{n} c_i^{2n-2} \end{vmatrix}. \tag{27}$$

(27)式表明,为了求出 $D(f)$,就需要计算

$$\sum_{i=1}^{n} c_i^k, \quad k=0,1,2,\cdots,2n-2. \tag{28}$$

由于 $f(x)$的 n 个复根 $c_1, c_2, \cdots, c_n$ 是未知的,因此必须想办法通过 $f(x)$的系数来计算 $\sum c_i^k$.注意到 $\sum c_i^k$ 是 $c_1, c_2, \cdots, c_n$ 的对称多项式,因此仍然想到运用对称多项式基本定理.为此我们考虑下列 n 元对称多项式

$$\begin{aligned} & s_k(x_1, x_2, \cdots, x_n) \\ = & x_1^k + x_2^k + \cdots + x_n^k, \quad k=0,1,2,\cdots \end{aligned} \tag{29}$$

这些 n 元对称多项式称为幂和.

据对称多项式基本定理,幂和 s_k 能表示成初等对称多项式的多项式.具体的表示方法可以用下述递推公式求出.

牛顿(Newton)公式[①]　当 $1 \le k \le n$ 时,

$$s_k - \sigma_1 s_{k-1} + \sigma_2 s_{k-2} + \cdots + (-1)^{k-1} \sigma_{k-1} s_1 + (-1)^k k\sigma_k = 0; \tag{30}$$

当 $k > n$ 时,

$$s_k - \sigma_1 s_{k-1} + \sigma_2 s_{k-2} + \cdots + (-1)^{n-1} \sigma_{n-1} s_{k-n+1} + (-1)^n \sigma_n s_{k-n} = 0. \tag{31}$$

利用牛顿公式,可以从 $s_{k-1}(c_1, \cdots, c_n), \cdots, s_1(c_1, \cdots, c_n)$以及 $\sigma_1(c_1, \cdots, c_n), \cdots, \sigma_n(c_1, \cdots, c_n)$计算出 $s_k(c_1, \cdots, c_n)$.再从(27)式就可以求出 $f(x)$的判别式 $D(f)$.

$f(x)$的判别式 $D(f)$也称为方程 $f(x)=0$ 的判别式.

注意:在上述讨论中,$f(x)$的首项系数为 1.如果 $f(x)$的首项系数为 a_n,则可以先对 $a_n^{-1} f(x)$运用上述方法求出它的判别式 $D(a_n^{-1} f)$,然后规定 $f(x)$的判别式为

$$D(f) := a_n^{2n-2} D(a_n^{-1} f). \tag{32}$$

例 3　求数域 K 上二次方程 $f(x) = x^2 + bx + c = 0$ 的判别式.

---

①　牛顿公式的证明可参看丘维声编著《高等代数下册》(高等教育出版社 1996 年 12 月第 1 版)第 125～127 页.

解　设 $f(x)$的复根是 c_1, c_2,则

$$\sigma_1(c_1, c_2) = -b, \quad \sigma_2(c_1, c_2) = c.$$

据牛顿公式得

$$s_1 = \sigma_1, \quad s_2 = \sigma_1 s_1 - 2\sigma_2 = \sigma_1^2 - 2\sigma_2,$$

于是
$$s_1(c_1, c_2) = -b, \quad s_2(c_1, c_2) = b^2 - 2c,$$

因此
$$D(f) = \begin{vmatrix} 2 & -b \\ -b & b^2 - 2c \end{vmatrix} = b^2 - 4c. \tag{33}$$

读者已从中学代数里知道 $b^2 - 4c$ 是 $x^2 + bx + c = 0$ 的判别式,现在知道了“判别式”一词来源在这里.

例 4　求数域 K 上不完全三次方程

$$f(x) = x^3 + ax + b = 0 \tag{34}$$

的判别式.

解　设 $f(x)$的复根是 c_1, c_2, c_3,则

$$\sigma_1(c_1, c_2, c_3) = 0, \quad \sigma_2(c_1, c_2, c_3) = a, \quad \sigma_3(c_1, c_2, c_3) = -b.$$

据牛顿公式得

$$s_1(c_1, c_2, c_3) = \sigma_1(c_1, c_2, c_3) = 0, \quad s_2(c_1, c_2, c_3) = -2a,$$

$$s_3(c_1, c_2, c_3) = -3b, \quad s_4(c_1, c_2, c_3) = 2a^2,$$

所以

$$D(f) = \begin{vmatrix} 3 & 0 & -2a \\ 0 & -2a & -3b \\ -2a & -3b & 2a^2 \end{vmatrix} = -4a^3 - 27b^2. \tag{35}$$

*如果考虑完全三次方程

$$f(x) = x^3 + a_1 x^2 + a_2 x + a_3 = 0, \tag{36}$$

则 $D(f)$具有比(35)更复杂的形式.但是我们可以借助下述办法把完全三次方程化为不完全三次方程:x 用 $x - \frac{a_1}{3}$代入,令

$$g(x) := f\left(x - \frac{a_1}{3}\right),$$

直接计算 $f\left(x - \frac{a_1}{3}\right)$可知,$g(x)$是不完全三次多项式($x^2$ 的系数为 0).如果我们求出了 $g(x)$的一个根 c_0,则

$$f\left(c_0 - \frac{a_1}{3}\right) = g(c_0) = 0,$$

从而 $c_0 - \frac{a_1}{3}$是 $f(x)$的一个根.因此把 $g(x)$的根弄清楚了,随之 $f(x)$的根也就

清楚了.所以不失一般性,我们总可以假定 x^2 的系数 $a_1=0$.

习 题 7.10

1. 设 $f(x_1,x_2,x_3)$是数域 K 上的一个三元多项式:

$$f(x_1,x_2,x_3)=x_1^3x_2^2+x_1^3x_3^2+x_1^2x_2^3+x_1^2x_3^3+x_2^3x_3^2+x_2^2x_3^3,$$

证明 $f(x_1,x_2,x_3)$是对称多项式.

2. 在 $K[x_1,x_2,x_3]$的含有项 $x_1^3x_2$ 的对称多项式中,写出项数最少的那个对称多项式.

3. 在 $K[x_1,x_2,x_3]$中,用初等对称多项式表出下列对称多项式:

(1) $x_1^3x_2+x_1x_2^3+x_1^3x_3+x_1x_3^3+x_2^3x_3+x_2x_3^3$;

(2) $x_1^4+x_2^4+x_3^4$;

(3) $(x_1x_2+x_3^2)(x_2x_3+x_1^2)(x_3x_1+x_2^2)$.

4. 在 $K[x_1,\cdots,x_n]$中,用初等对称多项式表出下列对称多项式($n\geq3$):

(1) $\sum x_1^3$;　　(2) $\sum x_1^2x_2^2x_3$.

5. 在 $K[x_1,x_2,x_3]$中,利用牛顿公式把幂和 s_2,s_3,s_4 用初等对称多项式表示.

6. 设 $f(x)$是实系数三次多项式,证明:当 $D(f)=0$时,$f(x)$在 C 中有重根,并且 $f(x)$在 C 中的根都是实数;当 $D(f)>0$ 时,$f(x)$有三个互不相同的实根;当 $D(f)<0$ 时,$f(x)$有一个实根,一对共轭虚根.

7. 设 $f(x)\in K[x]$,并且 $f(x)$的首项系数为 1,$a\in K$,$g(x)=(x-a)f(x)$,证明:

$$D(g)=D(f)f(a)^2.$$

8. 求 $f(x)=x^n+a\in K[x]$的判别式.

9. 求数域 K 上完全三次方程

$$f(x)=x^3+a_1x^2+a_2x+a_3=0$$

的判别式.

§11 有 限 域

我们已经知道,Z,$M_n(K)$,$K[x]$,$K[x_1,x_2,\cdots,x_n]$都是环.这些环都有无穷多个元素,称它们为无限环.有没有只有有限多个元素的环?

给定一个正整数 4,考虑 Z 的下列子集:

$$\bar{0}\overset{\text{def}}{=}\{4k\mid k\in Z\},\quad \bar{1}\overset{\text{def}}{=}\{4k+1\mid k\in Z\},$$

$$\bar{2}\overset{\text{def}}{=}\{4k+2\mid k\in Z\},\quad \bar{3}\overset{\text{def}}{=}\{4k+3\mid k\in Z\}.$$

称它们为模 4 剩余类.$\bar{2}$ 中的 2 称为代表,$\bar{2}$ 中任何一个整数都可作为代表.其余

类推.令

$$\mathbf{Z}_4 \overset{\text{def}}{=} \{\bar{0},\bar{1},\bar{2},\bar{3}\}. \tag{1}$$

在 $\mathbf{Z}_4$ 中规定加法和乘法运算如下:

$$\bar{i}+\bar{j} \overset{\text{def}}{=} \overline{i+j}, \quad \bar{i}\cdot\bar{j} \overset{\text{def}}{=} \overline{ij}. \tag{2}$$

需要说明(2)式的定义是合理的,即与剩余类的代表的选取无关.设 $\bar{i}=\bar{a}, \bar{j}=\bar{b}$,则

$$i-a=4k, \quad j-b=4l.$$

从而

$$(i+j)-(a+b)=(i-a)+(j-b)=4(k+l),$$

$$\begin{aligned} ij-ab &= ij-aj+aj-ab=(i-a)j+a(j-b) \\ &= 4kj+4la=4(kj+la). \end{aligned}$$

因此

$$\overline{i+j}=\overline{a+b}, \quad \overline{ij}=\overline{ab}.$$

这证明了(2)式的定义是合理的.例如

$$\bar{2}+\bar{3}=\bar{5}=\bar{1}, \quad \bar{2}\cdot\bar{3}=\bar{6}=\bar{2}, \quad \bar{2}\cdot\bar{2}=\bar{4}=\bar{0}.$$

容易看出,$\bar{0}$ 是 $\mathbf{Z}_4$ 对于加法的零元素,$\bar{1}$ 是 $\mathbf{Z}_4$ 对于乘法的单位元,并且 $\mathbf{Z}_4$ 是一个有单位元的交换环,它有非平凡的零因子 $\bar{2}$. $\mathbf{Z}_4$ 是只含 4 个元素的环.只含有限多个元素的环称为有限环.

一般地,给定一个正整数 m,考虑 $\mathbf{Z}$ 的子集

$$\bar{i} \overset{\text{def}}{=} \{km+i \mid k\in \mathbf{Z}\}, \quad i=0,1,2,\cdots,m-1,$$

称 $\bar{i}$ 是模 m 剩余类,i 称为代表,$\bar{i}$ 中任何一个整数都可以作为代表.容易看出,对于 $a\in \mathbf{Z}$,有

$$a\in \bar{i} \Longleftrightarrow m \mid a-i. \tag{3}$$

从(3)式受到启发,我们在 $\mathbf{Z}$ 上定义一个二元关系如下:

$$a\sim b \overset{\text{def}}{\Longleftrightarrow} m \mid a-b \tag{4}$$

容易验证,～是具有反身性、对称性和传递性.因此～是一个等价关系.称它为模 m 同余关系,并且把 $a\sim b$ 记作

$$a\equiv b \pmod{m}, \tag{5}$$

读作"a 与 b 模 m 同余".显然,在模 m 同余关系下的等价类就是模 m 剩余类.整数集 $\mathbf{Z}$ 对于模 m 同余关系的商集(它由所有等价类组成)记作 $\mathbf{Z}/m\mathbf{Z}$,或 $\mathbf{Z}/(m)$,或 $\mathbf{Z}_m$.显然有

$$\mathbf{Z}_m=\{\bar{0},\bar{1},\bar{2},\cdots,\overline{m-1}\}. \tag{6}$$

在 $\mathbf{Z}_m$ 中规定加法和乘法运算如下:

$$\bar{i}+\bar{j} \overset{\text{def}}{=} \overline{i+j}, \quad \bar{i}\,\bar{j} \overset{\text{def}}{=} \overline{ij}. \tag{7}$$

与前面类似可证明(7)式的定义是合理的.容易验证,Z_m 成为一个有单位元的交换环,称它为模 m 剩余类环.在 Z_m 中定义减法运算如下:

$$\bar{i}-\bar{j}\overset{\text{def}}{=}\bar{i}+(-\bar{j}).\tag{8}$$

例如,模 2 剩余类环 $Z_2=\{\bar{0},\bar{1}\}$.显然它的非零元 $\bar{1}$ 可逆.

又如,模 5 剩余类环 $Z_5=\{\bar{0},\bar{1},\bar{2},\bar{3},\bar{4}\}$.由于

$$\bar{2}\cdot\bar{3}=\bar{6}=\bar{1},\quad \bar{4}\cdot\bar{4}=\overline{16}=\bar{1},$$

因此 Z_5 的每个非零元都可逆.

从 Z_2,Z_5 等例子抽象出下述概念:

定义 1　设 F 是一个有单位元 $1(\neq 0)$ 的交换环,如果 F 的每一个非零元都可逆,则称 F 是一个域.

显然,数域都是域,它们含有无穷多个元素,称它们为无限域.

从上面的例子看到,Z_2,Z_5 都是域,只含有限多个元素的域称为有限域.

Z_4 不是域(因为 $\bar{2}$ 不可逆).观察 Z_4 与 Z_2,Z_5 的不同之处,可以猜想有下述结论:

定理 1　如果 p 是一个素数,则模 p 剩余类环 Z_p 是一个域.

证明　只要证 Z_p 的每一个非零元 $\bar{a}$ 可逆,这里 $0<a<p$.由于 p 是素数,且 $p\nmid a$,因此 $(p,a)=1$.于是存在 $u,v\in Z$,使得

$$ua+vp=1.$$

由此得出,$\bar{u}\,\bar{a}+\bar{v}\,\bar{p}=\bar{1}$,即 $\bar{u}\,\bar{a}=\bar{1}$.因此 $\bar{a}$ 可逆.从而 Z_p 是一个域.　　x

Z_p 称为模 p 剩余类域,它恰好含 p 个元素.在 Z_p 中,

$$p\bar{1}=\underbrace{\bar{1}+\bar{1}+\cdots+\bar{1}}_{p\text{个}}=\bar{p}=\bar{0}.\tag{9}$$

而当 $0<l<p$ 时,容易看出,$l\bar{1}\neq\bar{0}$.

在数域 K 中,对于任意正整数 n,都有 $n1\neq 0$.

定义 2　设 F 是一个域,它的单位元记作 e.如果存在一个素数 p,使得 $pe=0$,而对于 $0<l<p$ 都有 $le\neq 0$,则称域 F 的特征为 p;如果对一切正整数 n,都有 $ne\neq 0$,则称域 F 的特征为 0.把域 F 的特征记作 $\mathrm{char}\,F$.

定理 2　设 F 是一个域,则 F 的特征或者为 0,或者为一个素数.

证明　设 F 的特征不为 0,则存在一个正整数 n,使得 $ne=0$.设 n 是使 $ne=0$ 成立的最小正整数.假如 n 不是素数,则 $n=n_1n_2$,其中 $0<n_i<n$,$i=1,2$.于是

$$(n_1e)(n_2e)=(n_1n_2)e=ne=0.\tag{10}$$

如果 $n_1e\neq 0$,则 n_1e 可逆.在(10)式两边乘以 $(n_1e)^{-1}$ 得,$n_2e=0$,这与 n 的选取矛盾.如果 $n_1e=0$,也与 n 的选取矛盾.因此 n 必为素数.从而 F 的特征为素

数 n. x

推论 3 如果域 F 的特征为素数 p,则

$$ne = 0 \qquad p \mid n.$$

证明 充分性.设 $p \mid n$,则 $n = lp$.于是

$$ne = (lp)e = l(pe) = 0.$$

必要性.设 $ne = 0$.又设 $n = hp + r, 0 \leq r < p$,则

$$0 = ne = (hp + r)e = hpe + re = re.$$

由于 $\mathrm{char}\, F = p$,且 $r < p$.因此由上式得,$r = 0$.从而 $n = hp$,即 $p \mid n$. x

推论 4 设域 F 的特征为素数 p, $a \in F$ 且 $a \neq 0$,则 $na = 0$ 当且仅当 $p \mid n$.

证明 $na = 0 \qquad n(ea) = 0 \qquad (ne)a = 0$

$$ne = 0 \qquad p \mid n.$$ x

推论 4 告诉我们,在特征为素数 p 的域 F 中,要注意识别零元素:对于 $a \neq 0$,有 $pa = 0$.

类似于数域 K 上的一元多项式(或多元多项式),我们可以定义任一域 F 上的一元多项式(或多元多项式),并且得出域 F 上的一元多项式环 $F[x]$(或 n 元多项式环 $F[x_1, x_2, \cdots, x_n]$).不难看出,有关数域 K 上一元多项式环 $K[x]$ 的结论,只要在它的证明中没有用到这个域含有无穷多个元素,那么它对于任一域 F 上的一元多项式环 $F(x)$ 也成立.还需要注意,如果域 F 的特征为素数 p,则 F 的任一元素的 p 倍等于零.

例如,对于数域 K 上的两个一元多项式 $f(x)$ 与 $g(x)$,如果它们不相等,那么由它们分别确定的多项式函数 f 与 g 也不相等.这个结论的证明需要用到数域 K 有无穷多个元素.因此这个结论对于有限域上的多项式就不成立.譬如,在 $Z_3[x]$ 中,设 $f(x) = x^3 - x, g(x) = 0$.显然 $f(x) \neq g(x)$.但是由 $f(x)$ 确定的多项式函数 f 满足

$$f(\bar{0}) = \bar{0}, \quad f(\bar{1}) = \bar{1}^3 - \bar{1} = \bar{0}, \quad f(\bar{2}) = \bar{2}^3 - \bar{2} = \bar{0},$$

因此 f 是零函数.而 $g(x)$ 确定的函数也是零函数.

在任一域 F 上的一元多项式环 $F[x]$ 中,也有不可约多项式的概念和唯一因式分解定理;也有根与一次因式的关系,等等.

习 题 7.11

1. 证明:域 F 中没有非平凡的零因子,从而域一定是整环.
2. 写出模 7 剩余类域 Z_7 中每个非零元的逆元.

*3. 令

$$F = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \middle| a, b \in R \right\},$$

证明：F 对于矩阵的加法与乘法成为一个域，并且域 F 与复数域同构.

4. 令

$$F = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \middle| a, b \in Z_3 \right\},$$

证明：F 是一个有 9 个元素的域，并且 char F = 3.

5. 证明：在特征为 p 的域 F 中，下式成立：

$$(a+b)^p = a^p + b^p.$$

6. 写出 $Z_2[x]$ 中所有一次多项式和二次不可约多项式.

7. 设 $f(x) = x^5 - x^2 + 1 \in Z[x]$，判断 $f(x)$ 在 Q 上是否不可约.

8. 设 $f(x) = x^4 + 3x^3 + 3x^2 - 5 \in Z[x]$，判断 $f(x)$ 在 Q 上是否不可约.

第8章 线性空间

现实世界的空间形式中表现为“线性”的有直线、平面.研究直线和平面的工具是向量.空间中所有向量组成的集合有加法、数量乘法运算,它们满足加法交换律、结合律等8条运算法则.

现实世界的数量关系中“线性”问题可以用线性方程组来处理.研究线性方程组的工具是 n 维向量(即 n 元有序数组).数域 K 上所有 n 元有序数组组成的集合 K^n,有加法、数量乘法运算,它们满足加法交换律、结合律等8条运算法则.

矩阵在研究线性方程组中发挥了重要作用,因此矩阵也是研究线性问题的有力工具.矩阵有加法、数量乘法、乘法三种运算.数域 K 上所有 $s\times n$ 矩阵组成的集合中,矩阵的加法、数量乘法运算满足加法交换律、结合律等8条运算法则.

从上述受到启发,本章我们将建立一个数学模型:线性空间,研究线性空间的结构.它是研究客观世界中线性问题的重要理论.即使对于非线性问题,经过局部化后,就可以运用线性空间的理论,或者用线性空间的理论研究非线性问题的某一侧面.

§1 线性空间的结构

前面讲的几何空间中的向量,n 维向量,矩阵三个例子有共同之处:一个集合,一个数域,两种运算(加法与数量乘法),8条运算法则.由此抽象出线性空间的概念.

定义1 设 V 是一个非空集合,F 是一个域.在 V 上定义了一个代数运算:$(\alpha,\beta)\longmapsto\gamma$,叫做加法,把 γ 称为 α 与 β 的和,记作 $\gamma=\alpha+\beta$.在 F 与 V 之间定义了一个运算,即 $F\times V$ 到 V 的一个映射:$(k,\alpha)\longmapsto\delta$,叫做纯量乘法(当 F 为数域时,也叫数量乘法),把 δ 称为 k 与 α 的纯量乘积(当 F 为数域时,也叫数量乘积),记作 $\delta=k\alpha$.如果加法和数量乘法满足下述8条运算法则:对任意的 $\alpha,\beta,\gamma\in V$,任意的 $k,l\in F$,有

1° $\alpha+\beta=\beta+\alpha$ (加法交换律);

2° $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$ (加法结合律);

3° V 中有一个元素,记作 0,它使得

$$\alpha+0=\alpha, \forall\ \alpha\in V,$$

具有这个性质的元素 0 称为 V 的零元素;

4° 对于 $\alpha\in V$,存在 $\beta\in V$,使得

$$\alpha+\beta=0,$$

具有这个性质的元素 β 称为 α 的负元素;

5° $1\alpha=\alpha$;

6° $(kl)\alpha=k(l\alpha)$;

7° $(k+l)\alpha=k\alpha+l\alpha$;

8° $k(\alpha+\beta)=k\alpha+k\beta$,

则称 V 是域 F 上的一个线性空间.

借助几何语言,把线性空间的元素称为向量,线性空间又可称为向量空间. 域 F 上线性空间 V 的加法与纯量乘法运算统称为线性运算.

几何空间中以原点为起点的所有向量组成的集合,对于向量的加法与数量乘法运算,成为实数域 R 上的一个线性空间.

数域 K 上所有 n 元有序数组组成的集合 K^n,对于有序数组的加法与数量乘法运算,成为数域 K 上的一个线性空间. 在第 3 章中,我们把 K^n 称为数域 K 上的 n 维向量空间.

数域 K 上所有 $s\times n$ 矩阵组成的集合,对于矩阵的加法与数量乘法运算,成为数域 K 上的一个线性空间,记作 $M_{s\times n}(K)$.

下面我们再举一些线性空间的例子.

例 1　设 X 为实数域 R 的任一非空子集,定义域为 X 的所有实值函数组成的集合记作 R^X,它对于函数的加法,以及实数与函数的数量乘法,成为 R 上的一个线性空间.

例 2　设 X 为任意一个非空集合,F 是一个域,从 X 到 F 的每一个映射称为 X 上的一个 F 值函数. X 上的所有 F 值函数组成的集合记作 F^X. 在 F^X 中定义加法与纯量乘法运算如下:对于 $f,g\in F^X$, $k\in F$,规定

$$(f+g)(x)\overset{\text{def}}{=}f(x)+g(x), \forall\ x\in X,$$

$$(kf)(x)\overset{\text{def}}{=}k(f(x)), \forall\ x\in X,$$

容易验证它们满足加法交换律、结合律等 8 条运算法则. 因此 F^X 是 F 上的一个线性空间. F^X 的零元素是零函数,记作 0,即

$$0(x)=0, \forall\ x\in X.$$

例 3　数域 K 上所有一元多项式组成的集合 $K[x]$,它对于多项式的加法,以及 K 中元素与多项式的数量乘法,成为 K 上一个线性空间.

例 4　复数域 C 可以看成是实数域 R 上的一个线性空间,其加法是复数的

加法,其数量乘法是实数 a 与复数 z 相乘.

例 5　数域 K 可以看成是自身上的线性空间,其加法就是数域 K 中的加法,数量乘法就是 K 中的乘法.

上述例子表明,线性空间这一数学模型适用性很广.从现在开始,我们将从线性空间的定义出发,作逻辑推理,深入揭示线性空间的性质和结构,它们对于所有的具体的线性空间都成立.

设 V 是域 F 上的任一线性空间.

1. V 中零元素是唯一的.

证明　假设 $0_1, 0_2$ 是 V 中两个零元素,则

$$0_1 + 0_2 = 0_1,$$

$$0_1 + 0_2 = 0_2 + 0_1 = 0_2,$$

因此 $0_1 = 0_2$.　x

2. V 中每个元素 α 的负元素是唯一的.

证明　假设 β_1, β_2 都是 α 的负元素,则

$$(\beta_1 + \alpha) + \beta_2 = (\alpha + \beta_1) + \beta_2 = 0 + \beta_2 = \beta_2 + 0 = \beta_2,$$

$$\beta_1 + (\alpha + \beta_2) = \beta_1 + 0 = \beta_1.$$

根据加法结合律,得 $\beta_1 = \beta_2$.　x

今后把 α 的唯一的负元素记成 $-\alpha$.

利用负元素,可以在 V 中定义减法如下:对于 $\alpha, \beta \in V$,

$$\alpha - \beta \overset{\text{def}}{=} \alpha + (-\beta).$$

3. $0\alpha = 0, \forall \alpha \in V$.

证明　$0\alpha + \alpha = 0\alpha + 1\alpha = (0 + 1)\alpha = 1\alpha = \alpha$.

上式两边加上$(-\alpha)$,得

$$(0\alpha + \alpha) + (-\alpha) = \alpha + (-\alpha).$$

根据加法结合律和运算法则 4°,3°,得

$$0\alpha = 0.$$　x

4. $k0 = 0, \forall k \in K$.

证明　$$k0 + k0 = k(0 + 0) = k0.$$

上式两边加上 $-k0$,得

$$(k0 + k0) + (-k0) = k0 + (-k0).$$

根据加法结合律和运算法则 4°,3°,得

$$k0 = 0.$$　x

5. 如果 $k\alpha = 0$,那么 $k = 0$ 或 $\alpha = 0$.

证明　假设 $k \neq 0$,则

$$\alpha = 1\alpha = (k^{-1}k)\alpha = k^{-1}(k\alpha) = k^{-1}0 = 0.$$ x

6. $(-1)\alpha = -\alpha$, A$\alpha \in V$.

证明　$\alpha + (-1)\alpha = 1\alpha + (-1)\alpha = [1 + (-1)]\alpha = 0\alpha = 0$.

从 α 的负元素的唯一性,得 $(-1)\alpha = -\alpha$. x

设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是 V 中一个向量组,任给 F 中一组元素 $k_1, k_2, \cdots, k_s$,向量 $k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s$ 称为 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的一个线性组合,其中 $k_1, k_2, \cdots, k_s$ 称为系数.

对于 $\beta \in V$,如果有 F 中一组元素 $c_1, c_2, \cdots, c_s$,使得

$$\beta = c_1\alpha_1 + c_2\alpha_2 + \cdots + c_s\alpha_s,$$

则称 β 可以由 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性表出.

定义 2　V 中向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s\ (s \geq 1)$ 称为是线性相关的,如果有 F 中不全为零的元素 $k_1, k_2, \cdots, k_s$,使得

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0. \tag{1}$$

否则,向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 称为是线性无关的.即,如果从

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s = 0.$$

可以推出 $k_1 = k_2 = \cdots = k_s = 0$,则称向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是线性无关的.

我们把对应于 $s = 0$ 的空向量组定义成线性无关的,这可以为今后的讨论带来方便.

V 的任一子集称为向量集.对于有限向量集 W,如果给它的元素一种编号(即给它的元素排一个次序),写成 $\alpha_1, \alpha_2, \cdots, \alpha_s$,则称它是一个向量组.因此今后谈到向量组时,是指它有有限多个向量,并且排了一个次序.

V 的非空有限子集 W 称为是线性相关(线性无关)的,如果对于 W 中元素的一种编号(从而每一种编号)所得的向量组是线性相关(线性无关)的.

定义 3　设 W 是 V 的任一无限子集,如果 W 有一个有限子集是线性相关的,则称 W 是线性相关的;如果 W 的任何有限子集都是线性无关的,则称 W 是线性无关的.

例 6　单个向量 α 线性相关　　$\alpha = 0$.

例 7　包含零向量的向量集一定线性相关.

与本套教材上册第 3 章 §2 的方法类似,我们可以证明下述结论.这里我们要指出,本套教材上册第 1 章至第 6 章讲的数域 K 上的线性方程组的理论和数域 K 上的矩阵理论(包括行列式的理论)在把数域 K 换成任意域 F 后仍然成立.

命题 1　元素个数大于或等于 2 的向量集 W 线性相关当且仅当 W 中至少有一个向量可以由其余向量中的有限个线性表出.

命题2　设向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性无关，则向量 β 可以由 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出的充分必要条件是 $\alpha_1,\alpha_2,\cdots,\alpha_s,\beta$ 线性相关.

与本套教材上册第3章§3的方法类似，我们引出下述概念，并且可以证明下述结论.

定义4　设 W_1，W_2 都是 V 的非空子集，如果 W_1 中每一个向量都可以由 W_2 中有限多个向量线性表出，则称 W_1 可以由 W_2 线性表出. 如果 W_1 与 W_2 可以互相线性表出，则称 W_1 与 W_2 是等价的.

容易证明，线性表出有传递性，从而 V 中向量集的等价具有传递性. 显然，向量集的等价有反身性和对称性.

引理1　设向量组 $\beta_1,\beta_2,\cdots,\beta_r$ 可以由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出. 如果 $r>s$，那么向量组 $\beta_1,\beta_2,\cdots,\beta_r$ 线性相关.

推论3　设向量组 $\beta_1,\beta_2,\cdots,\beta_r$ 可以由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出. 如果 $\beta_1,\beta_2,\cdots,\beta_r$ 线性无关，则 $r\leq s$.

推论4　等价的线性无关的向量组所含向量的个数相等.

定义5　向量组(集)的一个部分组(子集)称为一个极大线性无关组(集)，如果这个部分组(子集)本身是线性无关的，但是从这个向量组(集)的其余向量(如果还有的话)中任取一个添进去，得到的新的部分组(子集)都线性相关.

推论5　向量组(集)与它的极大线性无关组(集)等价.

从推论5得，向量组(集)的任意两个极大线性无关组(集)等价. 从而有

推论6　向量组的任意两个极大线性无关组所含向量的个数相等.

定义6　向量组的一个极大线性无关组所含向量的个数称为这个向量组的秩.

全由零向量组成的向量组的秩为零.

向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的秩记作 $\mathrm{rank}\{\alpha_1,\alpha_2,\cdots,\alpha_s\}$.

命题7　向量组线性无关的充分必要条件是它的秩等于它所含向量的个数.

命题8　如果向量组(Ⅰ)可以由向量组(Ⅱ)线性表出，则(Ⅰ)的秩≤(Ⅱ)的秩.

推论9　等价的向量组有相同的秩.

几何空间的结构被它的一个基(3个不共面的向量)决定.

数域 K 上 n 元齐次线性方程组的解空间的结构被它的一个基(一个基础解系)决定.

数域 K 上 n 维向量空间 K^n 的任一非零子空间 U 的结构被 U 的一个基决定.

由这些受到启发,猜想域 F 上任一线性空间 V 也有基,且 V 的结构由它的一个基决定.

设 V 是域 F 上的任一线性空间.

定义 7　V 中的向量集 S 如果满足下述两个条件:

1°　向量集 S 是线性无关的;

2°　V 中每一个向量可以由 S 中有限多个向量线性表出,

则称 S 是 V 的一个基.

只含有零向量的线性空间的基为空集.

可以证明:域 F 上任一线性空间 V 都有基[①].

例 8　数域 K 上所有 $s\times n$ 矩阵形成的线性空间 $M_{s\times n}(K)$中,所有基本矩阵组成的子集

$$\{E_{11},E_{12},\cdots,E_{1n},\cdots,E_{s1},E_{s2},\cdots,E_{sn}\}$$

是 $M_{s\times n}(K)$的一个基.

证明　每个 $s\times n$ 矩阵 $A=(a_{ij})$可以表示成

$$A=\sum_{i=1}^{s}\sum_{j=1}^{n}a_{ij}E_{ij}.$$

假如 $\sum_{i=1}^{s}\sum_{j=1}^{n}k_{ij}E_{ij}=0$,则矩阵$(k_{ij})$是零矩阵,从而

$$k_{ij}=0,1\leq i\leq s,1\leq j\leq n.$$

因此$\{E_{11},E_{12},\cdots,E_{1n},\cdots,E_{s1},E_{s2},\cdots,E_{sn}\}$线性无关,从而它是 $M_{s\times n}(K)$的一个基.　x

例 9　数域 K 上所有一元多项式形成的线性空间 $K[x]$中,子集

$$S=\{1,x,x^2,\cdots,x^n,\cdots\}$$

是 $K[x]$的一个基.

证明　K 上每一个一元多项式 $f(x)$可以写成

$$f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n.$$

任取 S 的一个有限子集$\{x^{i_1},x^{i_2},\cdots,x^{i_m}\}$.假设

$$k_1x^{i_1}+k_2x^{i_2}+\cdots+k_mx^{i_m}=0,$$

则据一元多项式的定义得,$k_1=k_2=\cdots=k_m=0$,从而这个子集线性无关,因此 S 线性无关,于是 S 是 $K[x]$的一个基.　x

定义 8　V 称为有限维的,如果 V 有一个基包含有限多个向量;否则,V 称为无限维的.

---

①　证明可参看丘维声编著《高等代数(下册)》(高等教育出版社 1996 年 12 月第 1 版)第 176～177 页.

例 8 中的 $M_{s\times n}(K)$ 是有限维的,例 9 中的 $K[x]$ 是无限维的(容易看出,$K[x]$ 不可能有一个基是包含有限多个元素).

定理 10　如果 V 是有限维的,则 V 的任意两个基所含向量的个数相等.

证明　由定义 8 知,V 有一个基包含有限多个向量:$\alpha_1,\alpha_2,\cdots,\alpha_n$.设向量集 S 是 V 的另一个基.假如 S 包含的向量个数多于 n 个,则 S 中可取出 $n+1$ 个向量:$\beta_1,\beta_2,\cdots,\beta_{n+1}$.它们可以由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出.据引理 1 得,$\beta_1,\beta_2,\cdots,\beta_{n+1}$ 线性相关.这与 S 线性无关矛盾,因此 $|S|\leq n$.设 $S=\{\beta_1,\beta_2,\cdots,\beta_m\}$.据推论 4 得,$m=n$.　x

定义 9　设 V 是有限维的,则 V 的一个基所含向量的个数称为 V 的维数,记作 $\dim_F V$,简记作 $\dim V$.

只含零向量的线性空间的维数为 0.

由例 8 知道,$\dim M_{s\times n}(K)=sn$.

维数对于研究有限维线性空间的结构起着重要的作用.

命题 11　如果 $\dim V=n$,则 V 中任意 $n+1$ 个向量都线性相关.

证明　从定理 10 的证明过程可看出.　x

命题 12　如果 $\dim V=n$,则 V 中任意 n 个线性无关的向量都是 V 的一个基.

证明　在 V 中任取 n 个线性无关的向量 $\alpha_1,\alpha_2,\cdots,\alpha_n$.对于任意 $\beta\in V$,据命题 11 得,$\alpha_1,\alpha_2,\cdots,\alpha_n,\beta$ 线性相关.从而据命题 2 得,β 可以由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出.因此 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 的一个基.　x

基对于研究线性空间的结构起着重要作用.

命题 13　设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 的一个基,则 V 中每一个向量 α 可以唯一地表成 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的线性组合.

证明　由基的定义知道,α 可以由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出.假如有两种表出方式:

$$\alpha=a_1\alpha_1+a_2\alpha_2+\cdots+a_n\alpha_n,$$

$$\alpha=b_1\alpha_1+b_2\alpha_2+\cdots+b_n\alpha_n,$$

则得

$$0=(a_1-b_1)\alpha_1+(a_2-b_2)\alpha_2+\cdots+(a_n-b_n)\alpha_n.$$

由于 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关,因此

$$a_1-b_1=0,a_2-b_2=0,\cdots,a_n-b_n=0.$$

由此得出表出方式唯一.　x

我们把 α 由基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出的系数组成的 n 元有序组 $(a_1,a_2,\cdots,$

a_n)称为 α 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标.通常把向量 α 的坐标写成列向量的形式.

由上述看出,有限维线性空间 V 中给定一个基,则 V 中每一个向量都可以唯一地表示成这个基的线性组合,从而 V 的结构就很清楚了.

n 维线性空间 V 中给定两个基, V 中每一个向量分别在这两个基下的坐标有什么关系?

设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 与 $\beta_1,\beta_2,\cdots,\beta_n$ 是 V 的两个基. V 中向量 α 在这两个基下的坐标分别为

$$X=(x_1,x_2,\cdots,x_n)',\ Y=(y_1,y_2,\cdots,y_n)'.$$

为了求 X 与 Y 之间的关系,首先把这两个基之间的关系搞清楚.由于 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 的一个基,因此有

$$\begin{cases}\beta_1=a_{11}\alpha_1+a_{21}\alpha_2+\cdots+a_{n1}\alpha_n,\\ \beta_2=a_{12}\alpha_1+a_{22}\alpha_2+\cdots+a_{n2}\alpha_n,\\ \cdots\cdots\cdots\cdots\\ \beta_n=a_{1n}\alpha_1+a_{2n}\alpha_2+\cdots+a_{nn}\alpha_n.\end{cases}\tag{2}$$

为了使推导过程简洁,我们引进一种形式写法:

$$x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n\overset{\text{def}}{=}(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}.\tag{3}$$

进而可以把(2)式写成

$$(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{pmatrix}.\tag{4}$$

我们把(4)式右端的 n 级矩阵记作 A,称它是基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 到基 $\beta_1,\beta_2,\cdots,\beta_n$ 的过渡矩阵.于是(4)式可写成

$$(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A.\tag{5}$$

形式写法是模仿矩阵乘法的定义,因此类似于矩阵乘法的结合律、左(右)分配律、乘法与数量乘法的关系的证明方法,可以证明形式写法满足以下规则:

设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 与 $\beta_1,\beta_2,\cdots,\beta_n$ 是 V 中两个向量组, A,B 是域 F 上两个 n 级矩阵, $k\in F$,则

$$[(\alpha_1,\alpha_2,\cdots,\alpha_n)A]B=(\alpha_1,\alpha_2,\cdots,\alpha_n)(AB),\tag{6}$$

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)A+(\alpha_1,\alpha_2,\cdots,\alpha_n)B=(\alpha_1,\alpha_2,\cdots,\alpha_n)(A+B), \tag{7}$$

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)A+(\beta_1,\beta_2,\cdots,\beta_n)A=(\alpha_1+\beta_1,\alpha_2+\beta_2,\cdots,\alpha_n+\beta_n)A, \tag{8}$$

$$\begin{aligned}[k(\alpha_1,\alpha_2,\cdots,\alpha_n)]A &= (\alpha_1,\alpha_2,\cdots,\alpha_n)(kA)\\ &= k[(\alpha_1,\alpha_2,\cdots,\alpha_n)A],\end{aligned} \tag{9}$$

其中

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)+(\beta_1,\beta_2,\cdots,\beta_n) \stackrel{\text{def}}{=} (\alpha_1+\beta_1,\alpha_2+\beta_2,\cdots,\alpha_n+\beta_n), \tag{10}$$

$$k(\alpha_1,\alpha_2,\cdots,\alpha_n) \stackrel{\text{def}}{=} (k\alpha_1,k\alpha_2,\cdots,k\alpha_n). \tag{11}$$

命题 14 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 的一个基,且

$$(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A,$$

则 $\beta_1,\beta_2,\cdots,\beta_n$ 是 V 的一个基当且仅当 A 是可逆矩阵.

证明 由于 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性无关,并且有

$$\begin{aligned}k_1\beta_1+k_2\beta_2+\cdots+k_n\beta_n &= (\beta_1,\beta_2,\cdots,\beta_n)\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}\\ &= (\alpha_1,\alpha_2,\cdots,\alpha_n)A\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix},\end{aligned}$$

因此

$\beta_1,\beta_2,\cdots,\beta_n$ 是 V 的一个基

$\beta_1,\beta_2,\cdots,\beta_n$ 线性无关

从 $k_1\beta_1+k_2\beta_2+\cdots+k_n\beta_n=0$ 可推出 $k_1=k_2=\cdots=k_n=0$

从 $(\alpha_1,\alpha_2,\cdots,\alpha_n)A\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}=0$ 可推出 $\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}=0$

从 $A\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}=0$ 可推出 $\begin{pmatrix}k_1\\k_2\\\vdots\\k_n\end{pmatrix}=0$

齐次线性方程组 $AZ=0$ 只有零解

$$|A| \neq 0$$

A 是可逆矩阵.　　x

命题14的必要性表明,基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 到基 $\beta_1,\beta_2,\cdots,\beta_n$ 的过渡矩阵 A 是可逆矩阵.

现在可以给出向量 α 分别在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 与基 $\beta_1,\beta_2,\cdots,\beta_n$ 下的坐标 X,Y 之间的关系:由于

$$\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)X,\ \alpha=(\beta_1,\beta_2,\cdots,\beta_n)Y,$$

并且基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 到基 $\beta_1,\beta_2,\cdots,\beta_n$ 的过渡矩阵是 A,因此

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)X=(\beta_1,\beta_2,\cdots,\beta_n)Y$$
$$=(\alpha_1,\alpha_2,\cdots,\alpha_n)AY.$$

由于同一个向量由基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出的方式唯一,从上式得

$$X=AY, \tag{12}$$

从(12)式得

$$Y=A^{-1}X. \tag{13}$$

习　题　8.1

1. 判断下述集合对于所指的运算是否形成实数域 R 上的线性空间:

(1) $R[x]$ 中所有2次多项式组成的集合,对于多项式的加法与数量乘法;

(2) 所有正实数组成的集合 R^+,加法与数量乘法分别定义为

$$a \oplus b=ab,\ \forall\ a,b\in R^+,$$
$$k\odot a=a^k,\ \forall\ a\in R^+,k\in R;$$

(3) 区间 $[a,b]$ 上的所有连续函数组成的集合,记作 $C[a,b]$,对于函数的加法与数量乘法.

2. 判断实数域 R 上的线性空间 R^R 中的下列函数组是否线性无关?

(1) $1,\cos^2 x,\cos 2x$;

(2) $1,\cos x,\cos 2x,\cos 3x$;

(3) $1,\sin x,\cos x$

(4) $\sin x,\cos x,\sin^2 x,\cos^2 x$;

(5) $1,e^x,e^{2x},e^{3x},\cdots,e^{nx}$;

(6) $x^2,x|x|$.

3. 求第1题的(2)小题中线性空间的一个基和维数.

4. 把复数域 C 看成实数域 R 上的线性空间,求它的一个基和维数,以及每个复数 $z=a+bi$ 在这个基下的坐标.

5. 把数域 K 看成自身上的线性空间,求它的一个基和维数.

6. 说明数域 K 上所有 n 级对称矩阵组成的集合 V_1,对于矩阵的加法与数量乘法,形成 K 上一个线性空间,求 V_1 的一个基和维数.

7. 说明数域 K 上所有 n 级斜对称矩阵组成的集合 V_2,对于矩阵的加法与数量乘法,形成 K 上一个线性空间,求 V_2 的一个基和维数.

8. 说明数域 K 上所有 n 级上三角矩阵组成的集合 W,对于矩阵的加法与数量乘法,形成 K 上一个线性空间,求 W 的一个基和维数.

9. 在 K^3 中,设

$$\alpha_1=\begin{pmatrix}1\\0\\-1\end{pmatrix},\alpha_2=\begin{pmatrix}2\\1\\1\end{pmatrix},\alpha_3=\begin{pmatrix}1\\1\\1\end{pmatrix};$$

$$\beta_1=\begin{pmatrix}0\\1\\1\end{pmatrix},\beta_2=\begin{pmatrix}-1\\1\\0\end{pmatrix},\beta_3=\begin{pmatrix}1\\2\\1\end{pmatrix},$$

求基 $\alpha_1,\alpha_2,\alpha_3$ 到基 β_1,β_2,β_3 的过渡矩阵,并且求向量 $\alpha=(2,5,3)'$分别在这两个基下的坐标 X,Y.

10. 证明:在数域 K 上的 n 维线性空间 V 中,如果每一个向量都可由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出,则 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 的一个基.

11. 设 $X=\{x_1,x_2,\cdots,x_n\}$, F 是一个域.定义域为 X 的所有 F 值函数组成的集合记作 F^X,它是域 F 上的一个线性空间,求 F^X 的一个基和维数,并且求函数 f 在这个基下的坐标.

12. 设

$$V=\left\{\begin{pmatrix}x_1 & x_2+ix_3\\ x_2-ix_3 & -x_1\end{pmatrix}\middle|\, x_1,x_2,x_3\in R\right\}.$$

(1) 证明 V 对于矩阵的加法和数量乘法是实数域 R 上的一个线性空间;

(2) 求 V 的一个基和维数;

(3) 求 V 中元素 $\begin{pmatrix}x_1 & x_2+ix_3\\ x_2-ix_3 & -x_1\end{pmatrix}$ 在第(2)小题求出的一个基下的坐标.

§2　子空间及其交与和,子空间的直和

数域 K 上 n 元齐次线性方程组 $AX=0$ 的解空间 W 是 K^n 的子空间,意思是齐次线性方程组的解集 W 对于有序数组的加法与数量乘法封闭.

数域 K 上所有 n 级矩阵组成的集合记作 $M_n(K)$,它对于矩阵的加法与数量乘法形成 K 上一个线性空间. K 上所有 n 级对称矩阵组成的集合 V_1,对于矩阵的加法与数量乘法也形成 K 上一个线性空间(见习题 8.1 的第 6 题).显然 V_1 是 $M_n(K)$的子集,并且 V_1 的加法就是 $M_n(K)$的加法, V_1 的数量乘法就

是 $M_n(K)$的数量乘法.很自然地把 V_1 叫做 $M_n(K)$的一个子空间.

本节将介绍任意线性空间的子空间的概念、子空间的运算,以及研究如何利用子空间来刻画线性空间的结构.

定义 1　域 F 上线性空间 V 的一个非空子集 U 如果对于 V 的加法与纯量乘法也形成 F 上的线性空间,则称 U 是 V 的一个线性子空间,简称为子空间.

显然,$\{0\}$是 V 的一个子空间,称它为 V 的零子空间,也记作 0.

显然,V 是 V 的一个子空间.0 和 V 称为 V 的平凡子空间,其余的子空间称为非平凡子空间.

定理 1　域 F 上线性空间 V 的非空子集 U 是 V 的一个子空间当且仅当 U 对于 V 的加法与纯量乘法都封闭,即

1°　$u_1, u_2 \in U$　$u_1 + u_2 \in U$;

2°　$u \in U, k \in F$　$ku \in U$.

证明　必要性由定义 1 直接得出.现在证充分性.

由已知条件得,V 的加法与纯量乘法都是 U 的运算.由于 V 是线性空间,因此 U 的加法满足交换律、结合律;纯量乘法满足 5°,6°,7°,8°这 4 条法则.

由于 U 是非空集,因此有 $u \in U$.由已知条件得,$0u \in U$.由于 V 是线性空间,因此 $0u = 0$.从而 $0 \in U$,于是 V 的零元素是 U 的零元素.

任取 $\alpha \in U$,由已知条件得,$(-1)\alpha \in U$.由于 V 是线性空间,因此$(-1)\alpha = -\alpha$,从而 $-\alpha \in U$,于是 α 在 V 中的负元素 $-\alpha$ 也是 α 在 U 中的负元素.

综上述得,U 是 F 上一个线性空间.从而 U 是 V 的一个子空间.　x

例 1　数域 K 上所有次数小于 n 的一元多项式组成的集合记作 $K[x]_n$,证明 $K[x]_n$ 是 $K[x]$的一个子空间.

证明　显然 $K[x]_n$ 非空集.由于两个次数小于 n 的一元多项式的和的次数仍小于 n,且任一数 k 与一个次数小于 n 的一元多项式的乘积的次数仍小于 n,因此 $K[x]_n$ 对于多项式的加法与数量乘法都封闭,从而 $K[x]_n$ 是 $K[x]$的一个子空间.　x

命题 2　设 U 是域 F 上 n 维线性空间 V 的一个子空间,则

$$\dim U \leq \dim V.$$

证明　由于 n 维线性空间 V 中任意 $n+1$ 个向量都线性相关,因此 U 的一个基所含向量的个数一定小于或等于 n,从而

$$\dim U \leq \dim V.$$　x

命题 3　设 U 是域 F 上 n 维线性空间 V 的一个子空间,如果 $\dim U = \dim V$,则 $U = V$.

证明　由于 $\dim U = \dim V = n$,因此 U 的一个基 $\delta_1, \delta_2, \cdots, \delta_n$ 就是 V 的一

个基,从而 V 中任一向量 $\alpha = a_1\delta_1 + a_2\delta_2 + \cdots + a_n\delta_n \in U$,因此 $V \quad U$.显然 $U \quad V$,因此 $U = V$. x

命题 4　设 U 是域 F 上 n 维线性空间 V 的一个子空间,则 U 的一个基可以扩充成 V 的一个基.

证明　设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是 U 的一个基,则 $s \leq n$.如果 $s = n$,则 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 的一个基.下面设 $s < n$.此时 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 不是 V 的一个基,于是 V 中至少有一个向量 β_1 不能由 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性表出,从而 $\alpha_1, \alpha_2, \cdots, \alpha_s, \beta_1$ 线性无关.如果 $s + 1 = n$,则已得到 V 的一个基.如果 $s + 1 < n$,则同理有 $\beta_2 \in V$,使得 $\alpha_1, \alpha_2, \cdots, \alpha_s, \beta_1, \beta_2$ 线性无关.依次下去,到某一步,得到 n 个线性无关的向量 $\alpha_1, \alpha_2, \cdots, \alpha_s, \beta_1, \beta_2, \cdots, \beta_r$,其中 $s + r = n$,这就是 V 的一个基. x

如何构造域 F 上线性空间 V 的子空间？与我们在第 3 章 §1 中指出的方法类似,V 中给了向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$,由它们的所有线性组合组成的集合

$$\{k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s \mid k_1, k_2, \cdots, k_s \in F\}$$

是 V 的一个子空间,称它是由 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 生成的子空间,记作 $\langle \alpha_1, \alpha_2, \cdots, \alpha_s \rangle$.

与第 3 章 §4 的方法类似,可以证明下述结论:

定理 5　在域 F 上的线性空间 V 中,如果

$$U = \langle \alpha_1, \alpha_2, \cdots, \alpha_s \rangle,$$

则向量组 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 的一个极大线性无关组是 U 的一个基,从而

$$\dim U = \operatorname{rank}\{\alpha_1, \alpha_2, \cdots, \alpha_s\}.$$ x

从基的定义容易看出,如果 $\delta_1, \delta_2, \cdots, \delta_r$ 是 V 的子空间 U 的一个基,则 $U = \langle \delta_1, \delta_2, \cdots, \delta_r \rangle$.由此看出,$V$ 的任一有限维子空间都是由向量组生成的子空间.

几何空间 V 中,给了两个过原点 O 的平面 V_1, V_2,它们是 V 的子空间,如图 8 - 1 所示.从 V_1 与 V_2 能得到 V 的哪些子空间呢？

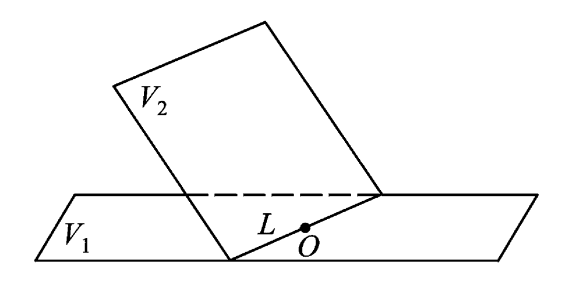


图 8 - 1

上述几何空间的例子中,平面 V_1 与 V_2 的交线 L 是 V 的一个子空间.一般地,我们有下述结论:

定理 6 设 V_1, V_2 都是域 F 上线性空间 V 的子空间,则 $V_1 \cap V_2$ 也是 V 的子空间.

证明 因为 $0 \in V_1 \cap V_2$,所以 $V_1 \cap V_2$ 非空集.设 $\alpha, \beta \in V_1 \cap V_2$,则 $\alpha, \beta \in V_i$, $i = 1, 2$.从而 $\alpha + \beta \in V_i$, $i = 1, 2$,因此 $\alpha + \beta \in V_1 \cap V_2$.同理可证,$V_1 \cap V_2$ 对于 V 的纯量乘法封闭.因此 $V_1 \cap V_2$ 是 V 的子空间. x

子空间的交适合交换律、结合律,即

$$V_1 \cap V_2 = V_2 \cap V_1, (V_1 \cap V_2) \cap V_3 = V_1 \cap (V_2 \cap V_3).$$

由结合律,我们可定义多个子空间的交: $\bigcap_{i=1}^{s} V_i$,它是 V 的一个子空间.

$V_1 \cup V_2$ 是不是 V 的一个子空间?从上述几何空间的例子看出,如果 $\alpha \in V_i$,且 $\alpha_i \notin L$, $i = 1, 2$,则虽然 $\alpha \in V_1 \cup V_2$, $i = 1, 2$,但是 $\alpha_1 + \alpha_2$ 可能不属于 $V_1 \cup V_2$,因此 $V_1 \cup V_2$ 不是 V 的子空间.如果我们想构造一个包含 $V_1 \cup V_2$ 的子空间,那么这个子空间应当包含 V_1 中任一向量 α_1 与 V_2 中任一向量 α_2 的和.由此受到启发,我们应当考虑下述集合

$$\{\alpha_1 + \alpha_2 \mid \alpha_1 \in V_1, \alpha_2 \in V_2\}.$$

定理 7 设 V_1, V_2 都是域 F 上线性空间 V 的子空间,则 V 的子集

$$\{\alpha_1 + \alpha_2 \mid \alpha_1 \in V_1, \alpha_2 \in V_2\}$$

是 V 的一个子空间,称它是 V_1 与 V_2 的和,记作 $V_1 + V_2$,即

$$V_1 + V_2 = \{\alpha_1 + \alpha_2 \mid \alpha_1 \in V_1, \alpha_2 \in V_2\}. \tag{1}$$

证明 由于 $0 = 0 + 0$,因此 $0 \in V_1 + V_2$.在 $V_1 + V_2$ 中任取两个向量 α, β,则 $\alpha = \alpha_1 + \alpha_2$, $\beta = \beta_1 + \beta_2$,其中 $\alpha_1, \beta_1 \in V_1$, $\alpha_2, \beta_2 \in V_2$.于是 $\alpha_1 + \beta_1 \in V_1$, $\alpha_2 + \beta_2 \in V_2$.因此

$$\begin{aligned} \alpha + \beta &= (\alpha_1 + \alpha_2) + (\beta_1 + \beta_2) \\ &= (\alpha_1 + \beta_1) + (\alpha_2 + \beta_2) \in V_1 + V_2. \end{aligned}$$

同理可证 $V_1 + V_2$ 对于 V 的纯量乘法封闭,因此 $V_1 + V_2$ 是 V 的一个子空间.

x

设 U 是 V 的子空间,且 $U \supseteq V_1 \cup V_2$,则 $U \supseteq V_1 + V_2$.这表明 $V_1 + V_2$ 是 V 中包含 $V_1 \cup V_2$ 的最小子空间.

从(1)式容易看出,子空间的和适合下述运算规则:

1° $V_1 + V_2 = V_2 + V_1$ (交换律);

2° $(V_1 + V_2) + V_3 = V_1 + (V_2 + V_3)$ (结合律).

由结合律,我们可以定义多个子空间的和:

$$V_1 + V_2 + \cdots + V_s$$

$$= \{\alpha_1 + \alpha_2 + \cdots + \alpha_s \mid \alpha_i \in V_i,\ i = 1, 2, \cdots, s\}, \quad (2)$$

它仍是 V 的一个子空间.

命题 8　设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 与 $\beta_1, \beta_2, \cdots, \beta_r$ 是域 F 上线性空间 V 的两个向量组,则

$$\langle \alpha_1, \alpha_2, \cdots, \alpha_s \rangle + \langle \beta_1, \beta_2, \cdots, \beta_r \rangle$$
$$= \langle \alpha_1, \alpha_2, \cdots, \alpha_s, \beta_1, \beta_2, \cdots, \beta_r \rangle.$$

证明　根据向量组生成的子空间的定义以及子空间的和的定义,得到

$$\langle \alpha_1, \alpha_2, \cdots, \alpha_s \rangle + \langle \beta_1, \beta_2, \cdots, \beta_r \rangle$$
$$= \{(k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s) + (l_1\beta_1 + l_2\beta_2 + \cdots + l_r\beta_r) \mid k_i, l_j \in F, 1 \leq i \leq s, 1 \leq j \leq r\}$$
$$= \langle \alpha_1, \alpha_2, \cdots, \alpha_s, \beta_1, \beta_2, \cdots, \beta_r \rangle.$$

x

如图 8-1 所示, $\dim V_1 = \dim V_2 = 2$, $\dim(V_1 \cap V_2) = 1$, 容易看出, $V_1 + V_2 = V$,因此 $\dim(V_1 + V_2) = 3$,于是有

$$\dim V_1 + \dim V_2 = \dim(V_1 + V_2) + \dim(V_1 \cap V_2).$$

对于任意 n 维线性空间,上式是否成立? 回答是肯定的.

定理 9(子空间的维数公式)　设 V_1, V_2 都是域 F 上线性空间 V 的有限维子空间,则 $V_1 \cap V_2$, $V_1 + V_2$ 也是有限维的子空间,并且

$$\dim V_1 + \dim V_2 = \dim(V_1 + V_2) + \dim(V_1 \cap V_2). \quad (3)$$

证明　由于 $V_1 \cap V_2 \quad V_1$,因此 $\dim(V_1 \cap V_2) \leq \dim V_1$. 设 V_1, V_2, $V_1 \cap V_2$ 的维数分别是 n_1, n_2, m. 在 $V_1 \cap V_2$ 中取一个基 $\alpha_1, \alpha_2, \cdots, \alpha_m$,把它分别扩充成 V_1, V_2 的一个基:

$$\alpha_1, \alpha_2, \cdots, \alpha_m, \beta_1, \cdots, \beta_{n_1 - m},$$
$$\alpha_1, \alpha_2, \cdots, \alpha_m, \gamma_1, \cdots, \gamma_{n_2 - m},$$

于是

$$V_1 + V_2$$
$$= \langle \alpha_1, \alpha_2, \cdots, \alpha_m, \beta_1, \cdots, \beta_{n_1 - m} \rangle + \langle \alpha_1, \alpha_2, \cdots, \alpha_m, \gamma_1, \cdots, \gamma_{n_2 - m} \rangle$$
$$= \langle \alpha_1, \alpha_2, \cdots, \alpha_m, \beta_1, \cdots, \beta_{n_1 - m}, \gamma_1, \cdots, \gamma_{n_2 - m} \rangle.$$

如果能证明向量组 $\alpha_1, \alpha_2, \cdots, \alpha_m, \beta_1, \cdots, \beta_{n_1 - m}, \gamma_1, \cdots, \gamma_{n_2 - m}$ 线性无关,则它是 $V_1 + V_2$ 的一个基,从而得出

$$\dim(V_1 + V_2) = m + (n_1 - m) + (n_2 - m) = n_1 + n_2 - m$$
$$= \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2).$$

假设有等式

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m + p_1\beta_1 + \cdots + p_{n_1-m}\beta_{n_1-m} + q_1\gamma_1 + \cdots + q_{n_2-m}\gamma_{n_2-m} = 0, \tag{4}$$

则

$$q_1\gamma_1 + \cdots + q_{n_2-m}\gamma_{n_2-m} = -k_1\alpha_1 - \cdots - k_m\alpha_m - p_1\beta_1 - \cdots - p_{n_1-m}\beta_{n_1-m}. \tag{5}$$

(5)式左边的向量属于 V_2,右边的向量属于 V_1,从而左边的向量属于 $V_1 \cap V_2$,因此它可由 $\alpha_1,\cdots,\alpha_m$ 线性表出:

$$q_1\gamma_1 + \cdots + q_{n_2-m}\gamma_{n_2-m} = l_1\alpha_1 + \cdots + l_m\alpha_m.$$

移项得

$$l_1\alpha_1 + l_2\alpha_2 + \cdots + l_m\alpha_m - q_1\gamma_1 - \cdots - q_{n_2-m}\gamma_{n_2-m} = 0. \tag{6}$$

由于 $\alpha_1,\alpha_2,\cdots,\alpha_m,\gamma_1,\cdots,\gamma_{n_2-m}$ 是 V_2 的一个基,因此从(6)式得出

$$l_1 = l_2 = \cdots = l_m = q_1 = \cdots = q_{n_2-m} = 0,$$

代入(4)式,得

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m + p_1\beta_1 + \cdots + p_{n_1-m}\beta_{n_1-m} = 0.$$

同理可得

$$k_1 = k_2 = \cdots = k_m = p_1 = \cdots = p_{n_1-m} = 0.$$

因此 $\alpha_1,\alpha_2,\cdots,\alpha_m,\beta_1,\cdots,\beta_{n_1-m},\gamma_1,\cdots,\gamma_{n_2-m}$ 线性无关. x

推论 10 设 V_1, V_2 都是域 F 上 n 维线性空间 V 的子空间,则

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 \qquad V_1 \cap V_2 = 0.$$ x

几何空间 V 中,V_1 是过原点的一个平面,V_2 是过原点的一条直线,且 V_2 不在平面 V_1 上.如图 8-2 所示,容易看出,$V_1 + V_2 = V$,且 $V_1 + V_2$ 中每个向量 α 能被唯一地表示成

$$\alpha = \alpha_1 + \alpha_2, \alpha_1 \in V_1, \alpha_2 \in V_2.$$

由此受到启发,引出下述概念:

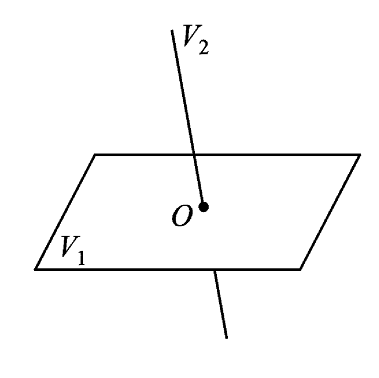


图 8-2

定义 2 设 V_1, V_2 是域 F 上线性空间 V 的子空间,如果 $V_1 + V_2$ 中每个向量 α 能唯一地表示成

$$\alpha = \alpha_1 + \alpha_2, \alpha_1 \in V_1, \alpha_2 \in V_2, \tag{7}$$

则称 $V_1 + V_2$ 是直和,记作 $V_1 \oplus V_2$.

定理 11 设 V_1, V_2 是域 F 上线性空间 V 的有限维子空间,则下列命题互相等价:

(i) $V_1 + V_2$ 是直和;

(ii) $V_1 + V_2$ 中零向量的表示法唯一;

(iii) $V_1 \cap V_2 = 0$;

(iv) $\dim(V_1 + V_2) = \dim V_1 + \dim V_2$;

(v) V_1 的一个基与 V_2 的一个基合起来是 $V_1 + V_2$ 的一个基.

证明 (i) (ii)是显然的.

(ii) (iii). 任取 $\alpha \in V_1 \cap V_2$,于是零向量可表成

$$0 = \alpha + (-\alpha), \alpha \in V_1, -\alpha \in V_2.$$

由已知条件得,$\alpha = 0$. 因此 $V_1 \cap V_2 = 0$.

(iii) (i). 任取 $\alpha \in V_1 + V_2$,假设 α 有两种表法:

$$\alpha = \alpha_1 + \alpha_2, \quad \alpha_1 \in V_1, \alpha_2 \in V_2,$$

$$\alpha = \beta_1 + \beta_2, \quad \beta_1 \in V_1, \beta_2 \in V_2,$$

则 $\alpha_1 + \alpha_2 = \beta_1 + \beta_2$,从而得到 $\alpha_1 - \beta_1 = \beta_2 - \alpha_2 \in V_1 \cap V_2$. 由于 $V_1 \cap V_2 = 0$,因此 $\alpha_1 = \beta_1, \alpha_2 = \beta_2$,从而 $V_1 + V_2$ 是直和.

(iii) (iv). 这由推论 10 立即得到.

(iv) (v). 设 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 是 V_1 的一个基,$\beta_1, \beta_2, \cdots, \beta_r$ 是 V_2 的一个基,则

$$V_1 + V_2 = \langle \alpha_1, \alpha_2, \cdots, \alpha_s \rangle + \langle \beta_1, \beta_2, \cdots, \beta_r \rangle$$
$$= \langle \alpha_1, \alpha_2, \cdots, \alpha_s, \beta_1, \beta_2, \cdots, \beta_r \rangle.$$

因为 $\dim(V_1 + V_2) = \dim V_1 + \dim V_2 = s + r$,且 $V_1 + V_2$ 的每一个向量可由 $\alpha_1, \alpha_2, \cdots, \alpha_s, \beta_1, \beta_2, \cdots, \beta_r$ 线性表出,所以 $\alpha_1, \alpha_2, \cdots, \alpha_s, \beta_1, \beta_2, \cdots, \beta_r$ 是 $V_1 + V_2$ 的一个基(参看习题 8.1 第 10 题).

(v) (iv) 是显然的. x

设 V_1, V_2 都是线性空间 V 的子空间,如果它们满足:

1° $V_1 + V_2 = V$;

2° $V_1 + V_2$ 是直和,

则称 V 是 V_1 与 V_2 的直和,记作 $V = V_1 \oplus V_2$. 此时称 V_1 是 V_2 的一个补空间,也称 V_2 是 V_1 的一个补空间.

命题 12　设 V 是域 F 上 n 维线性空间,则 V 的每一个子空间 U 都有补空间.

证明　U 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_m$,把它扩充成 V 的一个基:$\alpha_1,\alpha_2,\cdots,\alpha_m,\alpha_{m+1},\cdots,\alpha_n$,则

$$\begin{aligned}V &= \langle\alpha_1,\alpha_2,\cdots,\alpha_m,\alpha_{m+1},\cdots,\alpha_n\rangle \\ &= \langle\alpha_1,\alpha_2,\cdots,\alpha_m\rangle + \langle\alpha_{m+1},\cdots,\alpha_n\rangle = U + W,\end{aligned}$$

其中 $W=\langle\alpha_{m+1},\cdots,\alpha_n\rangle$.由于 U 的一个基与 W 的一个基合起来是 V 的一个基,因此 $U+W$ 是直和,从而

$$V = U \oplus W.$$

于是 W 是 U 的一个补空间.　□

例 2　设 $V=M_n(K)$,其中 K 是数域.分别用 V_1,V_2 表示 K 上所有 n 级对称、斜对称矩阵组成的子空间,证明:

$$V = V_1 \oplus V_2.$$

证明　第一步,证明 $V_1+V_2=V$.显然 $V_1+V_2 \subseteq V$.关键要证 $V \subseteq V_1+V_2$.任取 $A\in V=M_n(K)$,有

$$A = \frac{A+A'}{2} + \frac{A-A'}{2}. \tag{8}$$

容易验证 $\frac{A+A'}{2}$ 是对称矩阵,$\frac{A-A'}{2}$ 是斜对称矩阵,因此从(8)式得,$A\in V_1+V_2$.从而 $V \subseteq V_1+V_2$.因此 $V=V_1+V_2$.

第二步,证明 V_1+V_2 是直和,为此只要证 $V_1\cap V_2=0$.任取 $B\in V_1\cap V_2$,则 $B'=B$,且 $B'=-B$,于是 $B=-B$,即 $2B=0$,从而 $B=0$.因此 $V_1\cap V_2=0$.

综上所述,$V=V_1\oplus V_2$.　□

子空间的直和的概念可以推广到多个子空间的情形.

定义 3　设 $V_1,V_2,\cdots,V_s$ 都是域 F 上线性空间 V 的子空间,如果 $V_1+V_2+\cdots+V_s$ 中每一个向量 α 可唯一地表示成

$$\alpha=\alpha_1+\alpha_2+\cdots+\alpha_s,\ \alpha_i\in V_i,\ i=1,2,\cdots,s, \tag{9}$$

则称 $V_1+V_2+\cdots+V_s$ 是直和,记作 $V_1\oplus V_2\oplus\cdots\oplus V_s$ 或 $\bigoplus_{i=1}^{s} V_i$.

不难证明下述结论:

定理 13　设 $V_1,V_2,\cdots,V_s$ 都是域 F 上线性空间 V 的有限维子空间,则下列命题互相等价:

(i) $V_1+V_2+\cdots+V_s$ 是直和;

(ii) $V_1+V_2+\cdots+V_s$ 中零向量的表法唯一;

(iii) $V_i \cap \left(\sum_{j \neq i} V_j \right) = 0,\ i = 1, 2, \cdots, s$;

(iv) $\dim(V_1 + V_2 + \cdots + V_s) = \dim V_1 + \dim V_2 + \cdots + \dim V_s$;

(v) $V_i(i = 1, 2, \cdots, s)$的一个基,合起来是 $V_1 + V_2 + \cdots + V_s$的一个基.

证明的途径是:

(i)　(ii)　(iii)　(iv)　(v)

(i)　　(ii)

例 3　设 $V = K^4$, $V_1 = \langle \alpha_1, \alpha_2, \alpha_3 \rangle$, $V_2 = \langle \beta_1, \beta_2 \rangle$,其中

$$\alpha_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \alpha_2 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \alpha_3 = \begin{pmatrix} 0 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \beta_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \beta_2 = \begin{pmatrix} 1 \\ -1 \\ 3 \\ 7 \end{pmatrix},$$

分别求 $V_1 + V_2$, $V_1 \cap V_2$的一个基和维数.

解　因为

$V_1 + V_2 = \langle \alpha_1, \alpha_2, \alpha_3 \rangle + \langle \beta_1, \beta_2 \rangle = \langle \alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2 \rangle$,所以向量组 $\alpha_1, \alpha_2, \alpha_3$, β_1, β_2的一个极大线性无关组就是 $V_1 + V_2$的一个基,这个向量组的秩就是 $V_1 + V_2$的维数,为此令 $A = (\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2)$,对 A 作一系列初等行变换,化成简化行阶梯形矩阵:

$$A = \begin{pmatrix} 1 & -1 & 0 & 2 & 1 \\ 2 & 1 & 3 & -1 & -1 \\ 1 & 1 & 2 & 0 & 3 \\ 0 & 1 & 1 & 1 & 7 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{10}$$

由此得出, $\alpha_1, \alpha_2, \beta_1$是 $V_1 + V_2$的一个基, $\dim(V_1 + V_2) = 3$.同时也知道, β_2可由 $\alpha_1, \alpha_2, \beta_1$线性表出,其系数应当是线性方程组

$$x_1\alpha_1 + x_2\alpha_2 + x_3\beta_1 = \beta_2$$

的解.而从上述 A 及其简化行阶梯形矩阵的第 1,2,4,5 列可以看出,此方程组的解是(-1,4,3).因此

$$\beta_2 = -\alpha_1 + 4\alpha_2 + 3\beta_1,$$

从而
$$\alpha_1 - 4\alpha_2 = 3\beta_1 - \beta_2 \in V_1 \cap V_2.$$

因为
$$\dim(V_1 \cap V_2) = \dim V_1 + \dim V_2 - \dim(V_1 + V_2),$$

而从(10)可以看出, $\dim V_1 = 2$, $\dim V_2 = 2$,所以

$$\dim(V_1 \cap V_2) = 2 + 2 - 3 = 1,$$

于是 $\alpha_1 - 4\alpha_2 = (5, -2, -3, -4)'$是 $V_1 \cap V_2$的一个基.

例 3 的求解过程表明,只要把矩阵 A 经过一系列初等行变换化成简化行阶

梯形矩阵,那么我们所需要的一切信息都可由此得到.

习　题　8.2

1. 判断数域 K 上下列 n 元方程的解集是否为 K^n 的子空间:

(1) $a_1x_1+a_2x_2+\cdots+a_nx_n=0$;

(2) $a_1x_1+a_2x_2+\cdots+a_nx_n=1$;

(3) $x_1^2+x_2^2+x_3^2+\cdots+x_{n-1}^2-x_n^2=0$.

2. 设 A 是数域 K 上的一个 n 级矩阵.证明:数域 K 上所有与 A 可交换的矩阵组成的集合是 $M_n(K)$ 的一个子空间,把它记作 $C(A)$.

3. 设 $A=\text{diag}\{a_1,a_2,\cdots,a_n\}$,其中 $a_1,a_2,\cdots,a_n$ 是数域 K 中两两不同的数,求 $C(A)$ 的一个基和维数.

4. 设数域 K 上的 3 级矩阵

$$A=\begin{pmatrix}0&0&1\\1&0&0\\4&-2&1\end{pmatrix},$$

求 $C(A)$ 的维数和一个基.

5. 在域 F 上的线性空间 V 中,如果

$$k_1\alpha+k_2\beta+k_3\gamma=0,\text{且 } k_1k_2\neq0,$$

证明:$\langle\alpha,\gamma\rangle=\langle\beta,\gamma\rangle$.

6. 在 K^4 中(K 是数域),求向量组 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 生成的子空间的维数和一个基.设

$$\alpha_1=(1,-3,2,-1)',\qquad \alpha_2=(-2,1,5,3)',$$
$$\alpha_3=(4,-3,7,1)',\qquad \alpha_4=(-1,-11,8,-3)'.$$

7. 设 V 是域 F 上一个 n 维线性空间,$\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 的一个基,$\beta_1,\beta_2,\cdots,\beta_s$ 是 V 的一个向量组,并且

$$(\beta_1,\beta_2,\cdots,\beta_s)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A.$$

证明:$\langle\beta_1,\beta_2,\cdots,\beta_s\rangle$ 的维数等于 $n\times s$ 矩阵 A 的秩.

*8. 设实二次型

$$f(x_1,x_2,\cdots,x_n)=x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_n^2$$

的符号差 $s>0$,证明:方程 $f(x_1,\cdots,x_n)=0$ 的解集 W 中有一个子集 W_1 是 R^n 的一个线性子空间,且 $\dim W_1=n-p$.

9. 设 V_1,V_2 是域 F 上线性空间 V 的两个真子空间(即 $V_i\neq V$),证明:

$$V_1\cup V_2\neq V.$$

*10. 设 $V_1,V_2,\cdots,V_s$ 是域 F 上线性空间 V 的 s 个真子空间,证明:如果 $\text{char}\,F=0$,则 V 中至少有一个向量不属于 $V_1,V_2,\cdots,V_s$ 中任何一个.

11. 设 $V=K^4$,$V_1=\langle\alpha_1,\alpha_2\rangle$,$V_2=\langle\beta_1,\beta_2\rangle$,其中

$$\alpha_1=\begin{pmatrix}1\\-1\\0\\1\end{pmatrix},\alpha_2=\begin{pmatrix}-2\\3\\1\\-3\end{pmatrix},\beta_1=\begin{pmatrix}1\\2\\0\\-2\end{pmatrix},\beta_2=\begin{pmatrix}1\\3\\1\\-3\end{pmatrix},$$

分别求 V_1+V_2, $V_1\cap V_2$的一个基和维数.

12. 设 $V=K^4$, $V_1=\langle\alpha_1,\alpha_2,\alpha_3\rangle$, $V_2=\langle\beta_1,\beta_2\rangle$,其中

$$\alpha_1=\begin{pmatrix}1\\1\\-1\\2\end{pmatrix},\alpha_2=\begin{pmatrix}2\\-1\\3\\0\end{pmatrix},\alpha_3=\begin{pmatrix}0\\-3\\5\\-4\end{pmatrix},\beta_1=\begin{pmatrix}1\\2\\2\\1\end{pmatrix},\beta_2=\begin{pmatrix}4\\-3\\3\\1\end{pmatrix},$$

分别求 V_1+V_2, $V_1\cap V_2$的一个基和维数.

13. 设 $V=K^n$, K 是数域,齐次线性方程组

$$x_1+x_2+\cdots+x_n=0$$

的解空间记作 V_1;齐次线性方程组

$$\begin{aligned}x_1-x_2&=0,\\x_1-x_3&=0,\\&\cdots\cdots\\x_1-x_n&=0\end{aligned}$$

的解空间记作 V_2.证明: $V=V_1\oplus V_2$.

14. 证明:数域 K 上每一个 n 维线性空间 V 都可以表示成 n 个一维子空间的直和.

15. 用 $M_n^0(K)$表示 $M_n(K)$中迹为零的矩阵组成的集合, K 是数域.

(1) 证明: $M_n^0(K)$是线性空间 $M_n(K)$的一个子空间;

(2) 证明: $M_n(K)=\langle I\rangle\oplus M_n^0(K)$.

*16. 设 A 是数域 K 上的 n 级矩阵, $\lambda_1,\lambda_2,\cdots,\lambda_s$是 A 的全部不同的特征值,用 V_{λ_j} 表示 A 的属于λ_j的特征子空间.证明: A 可对角化的充分必要条件是

$$K^n=V_{\lambda_1}\oplus V_{\lambda_2}\oplus\cdots\oplus V_{\lambda_s}.$$

17. 设 K 是数域, $f_1(x),f_2(x)\in K[x]$, $A\in M_n(K)$.令 $f(x)=f_1(x)f_2(x)$.用 W, W_1, W_2分别表示齐次线性方程组

$$f(A)X=0,\quad f_1(A)X=0,\quad f_2(A)X=0$$

的解空间.

(1) 证明: W_1, W_2都是 W 的子空间;

(2) 证明:如果$(f_1(x),f_2(x))=1$,则 $W=W_1\oplus W_2$.

§3　线性空间的同构

域 F 上 n 维线性空间 V 与域 F 上 n 元有序组组成的线性空间 F^n 非常相

像.例如,对于 F^n 中向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 生成的子空间 $U=\langle\alpha_1,\alpha_2,\cdots,\alpha_s\rangle$,向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 的一个极大线性无关组是 U 的一个基,$\dim U$ 等于 $\mathrm{rank}\{\alpha_1,\alpha_2,\cdots,\alpha_s\}$.对于 V 中向量组生成的子空间也有同样的结论.

为什么域 F 上 n 维线性空间 V 与 F^n 这样相像?本节就来确切地阐述这种现象的实质.

定义 1 设 V 与 V' 都是域 F 上的线性空间,如果有 V 到 V' 的一个双射 σ,使得对于任意 $\alpha,\beta\in V$, $k\in F$,有

$$\sigma(\alpha+\beta)=\sigma(\alpha)+\sigma(\beta), \tag{1}$$

$$\sigma(k\alpha)=k\sigma(\alpha), \tag{2}$$

那么称 σ 是 V 到 V' 的一个同构映射(简称为同构).如果 V 到 V' 有一个同构映射,则称 V 与 V' 是同构的,记作

$$V\cong V'.$$

域 F 上线性空间 V 到 V' 的一个同构映射 σ 具有下列性质.

性质 1 $\sigma(0)$ 是 V' 的零元素 $0'$.

证明 因为 $0\alpha=0$,所以

$$\sigma(0)=\sigma(0\alpha)=0\sigma(\alpha)=0'. \qquad \blacksquare$$

性质 2 对于任意 $\alpha\in V$,有 $\sigma(-\alpha)=-\sigma(\alpha)$.

证明 $\sigma(-\alpha)=\sigma((-1)\alpha)=(-1)\sigma(\alpha)=-\sigma(\alpha)$. $\blacksquare$

性质 3 对于 V 中任一向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$, F 中任意一组元素 $k_1,k_2,\cdots,k_s$,有

$$\sigma(k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s)=k_1\sigma(\alpha_1)+k_2\sigma(\alpha_2)+\cdots+k_s\sigma(\alpha_s).$$

证明 由定义 1 即得. $\blacksquare$

性质 4 V 中向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性相关当且仅当 $\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_s)$ 是 V' 中线性相关的向量组.

证明 因为 σ 是 V 到 V' 的一个单射,所以如果 $\sigma(\alpha)=\sigma(\beta)$,则 $\alpha=\beta$.于是有

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s=0$$
$$\Leftrightarrow \sigma(k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s)=\sigma(0)$$
$$\Leftrightarrow k_1\sigma(\alpha_1)+k_2\sigma(\alpha_2)+\cdots+k_s\sigma(\alpha_s)=0',$$

从而 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性相关当且仅当 $\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_s)$ 线性相关. $\blacksquare$

性质 5 如果 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 的一个基,则 $\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)$ 是 V' 的一个基. $\blacksquare$

证明 据性质 4 得, $\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n)$ 是 V' 的一个线性无关的向量组.任取 $\beta\in V'$,由于 σ 是 V 到 V' 的一个满射,因此存在 $\alpha\in V$,使得 $\sigma(\alpha)=\beta$.

设

$$\alpha = a_1\alpha_1 + a_2\alpha_2 + \cdots + a_n\alpha_n,$$

则
$$\beta = \sigma(\alpha) = a_1\sigma(\alpha_1) + a_2\sigma(\alpha_2) + \cdots + a_n\sigma(\alpha_n),$$

因此 $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_n)$ 是 V' 的一个基. ■

定理 1 域 F 上两个有限维线性空间同构的充分必要条件是它们的维数相同.

证明 设 V 与 V' 都是域 F 上有限维线性空间.

必要性从性质 5 立即得出.

充分性. 设 $\dim V = \dim V' = n$. 在 V, V' 中各取一个基: $\alpha_1, \alpha_2, \cdots, \alpha_n$; $\gamma_1, \gamma_2, \cdots, \gamma_n$. 令

$$\sigma: V \to V'$$

$$\alpha = \sum_{i=1}^{n} a_i\alpha_i \longmapsto \sum_{i=1}^{n} a_i\gamma_i. \tag{3}$$

从(3)式看出, V 中每一个向量 α 都有 V' 中唯一的向量与 α 对应. 由于 $\gamma_1, \gamma_2, \cdots, \gamma_n$ 是 V' 的一个基, 因此 V 中不同的向量对应于 V' 中不同的向量; 并且 V' 中每一个向量 $\delta = \sum_{i=1}^{n} b_i\gamma_i$, 都有 V 中向量 $\beta = \sum_{i=1}^{n} b_i\alpha_i$ 对应于 δ. 因此 σ 是 V 到 V' 的一个双射. 设 $\alpha = \sum_{i=1}^{n} a_i\alpha_i$, $\beta = \sum_{i=1}^{n} b_i\alpha_i$, $k \in F$, 则

$$\begin{aligned}
\sigma(\alpha + \beta) &= \sigma\left(\sum_{i=1}^{n}(a_i + b_i)\alpha_i\right) = \sum_{i=1}^{n}(a_i + b_i)\gamma_i \\
&= \sum_{i=1}^{n} a_i\gamma_i + \sum_{i=1}^{n} b_i\gamma_i \\
&= \sigma(\alpha) + \sigma(\beta),
\end{aligned}$$

$$\begin{aligned}
\sigma(k\alpha) &= \sigma\left(\sum_{i=1}^{n}(ka_i)\alpha_i\right) = \sum_{i=1}^{n}(ka_i)\gamma_i \\
&= k\sum_{i=1}^{n} a_i\gamma_i = k\sigma(\alpha),
\end{aligned}$$

因此 σ 是 V 到 V' 的一个同构映射, 从而 $V \cong V'$. ■

从定理 1 立即得出, 域 F 上任一 n 维线性空间 V 都与 F^n 同构, 并且可以如下建立 V 到 F^n 的一个同构映射: 在 V 中取一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$; F^n 中取标准基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$. 令

$$\sigma: V \to F^n$$

$$\alpha = \sum_{i=1}^{n} a_i\alpha_i \longmapsto \sum_{i=1}^{n} a_i\varepsilon_i = (a_1, a_2, \cdots, a_n)', \tag{4}$$

即把 V 中每一个向量 α 对应到它在 V 的一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标 $(a_1, a_2, \cdots, a_n)'$,这就是 V 到 F^n 的一个同构映射.正是因为域 F 上任一 n 维线性空间 V 与 F^n 同构,所以 V 与 F^n 才这么相像,它们虽然元素不同,但是有关线性运算的性质却完全一样.从而我们可以利用 F^n 的性质来研究 F 上 n 维线性空间的性质.

命题 2　设 V 是域 F 上 n 维线性空间,U 是 V 的一个子空间.V 中取一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$, σ 把 V 中每一个向量 α 对应到它在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标.令

$$\sigma(U) \overset{\text{def}}{=} \{\sigma(\alpha) \mid \alpha \in U\},$$

则 $\sigma(U)$ 是 F^n 的一个子空间,且 $\dim U = \dim \sigma(U)$.

证明　显然 $\sigma(U)$ 是非空集.任取 $\alpha, \beta \in U$,有 $\alpha + \beta \in U$, $k\alpha \in U$.由于 σ 是 V 到 F^n 的一个同构映射,因此

$$\sigma(\alpha) + \sigma(\beta) = \sigma(\alpha + \beta) \in \sigma(U),$$

$$k\sigma(\alpha) = \sigma(k\alpha) \in \sigma(U),$$

因此 $\sigma(U)$ 是 F^n 的一个子空间.

由于 U 与 $\sigma(U)$ 都是域 F 上有限维线性空间,且 σ 限制到 U 上是 U 到 $\sigma(U)$ 的一个同构映射,因此据定理 1 的必要性得

$$\dim U = \dim \sigma(U).$$ x

例 1　设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是域 F 上线性空间 V 的一个基, $\beta_1, \beta_2, \cdots, \beta_s$ 是 V 的一个向量组,并且

$$(\beta_1, \beta_2, \cdots, \beta_s) = (\alpha_1, \alpha_2, \cdots, \alpha_n) A, \tag{5}$$

其中 A 是一个 $n \times s$ 矩阵.证明:

$$\dim \langle \beta_1, \beta_2, \cdots, \beta_s \rangle = \operatorname{rank}(A). \tag{6}$$

证明　用 σ 表示 V 到 F^n 的一个同构映射,它把 $\alpha \in V$ 映到 α 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标,设 A 的列向量组是: $A_1, A_2, \cdots, A_s$,从(5)式看出, $\sigma(\beta_j) = A_j$, $j = 1, 2, \cdots, s$.据命题 2 得,

$$\begin{aligned} \dim \langle \beta_1, \beta_2, \cdots, \beta_s \rangle &= \dim \sigma \langle \beta_1, \beta_2, \cdots, \beta_s \rangle \\ &= \dim \langle \sigma(\beta_1), \sigma(\beta_2), \cdots, \sigma(\beta_s) \rangle = \dim \langle A_1, A_2, \cdots, A_s \rangle \\ &= \operatorname{rank}(A). \end{aligned}$$ x

同构是域 F 上线性空间之间的一个关系,它具有反身性(因为 V 的恒等映射是 V 到 V 的一个同构映射),对称性和传递性(从下面的命题 3 得出).从而同构关系是一个等价关系,等价类称为同构类.

定理 1 表明,对于域 F 上所有有限维线性空间组成的集合 S 来说,维数为 0 的线性空间(即,$\{0\}$)恰好组成一个同构类,所有 1 维线性空间恰好组成一个同

构类,所有2维线性空间恰好组成一个同构类,……即维数完全决定了同构类.于是域 F 上有限维线性空间的同构类与非负整数之间有一个一一对应.由此看出,有限维线性空间的结构是如此简单!

命题3　域 F 上线性空间之间的一个同构映射的逆映射也是同构映射;两个同构映射的乘积还是同构映射.

证明　设 σ 是域 F 上线性空间 V 到 V' 的一个同构映射.显然 σ^{-1} 是 V' 到 V 的一个双射.任取 $\alpha', \beta' \in V'$,则存在 $\alpha, \beta \in V$,使得 $\alpha' = \sigma(\alpha), \beta' = \sigma(\beta)$.从而 $\sigma^{-1}(\alpha') = \alpha, \sigma^{-1}(\beta') = \beta$.于是

$$\begin{aligned}\sigma^{-1}(\alpha' + \beta') &= \sigma^{-1}(\sigma(\alpha) + \sigma(\beta)) = \sigma^{-1}[\sigma(\alpha + \beta)] \\ &= (\sigma^{-1}\sigma)(\alpha + \beta) = 1_V(\alpha + \beta) \\ &= \alpha + \beta = \sigma^{-1}(\alpha') + \sigma^{-1}(\beta');\end{aligned}$$

$$\sigma^{-1}(k\alpha') = \sigma^{-1}(k\sigma(\alpha)) = \sigma^{-1}(\sigma(k\alpha)) = \sigma^{-1}\sigma(k\alpha) = k\alpha = k\sigma^{-1}(\alpha').$$

因此 σ^{-1} 是 V' 到 V 的一个同构映射.

设 σ 和 τ 分别是域 F 上线性空间 V 到 V' 与 V' 到 V'' 的一个同构映射,则 $\tau\sigma$ 是 V 到 V'' 的一个双射,任取 $\alpha, \beta \in V$, $k \in F$,有

$$(\tau\sigma)(\alpha + \beta) = \tau(\sigma(\alpha + \beta)) = \tau[\sigma(\alpha) + \sigma(\beta)] = \tau\sigma(\alpha) + \tau\sigma(\beta),$$

$$(\tau\sigma)(k\alpha) = \tau(\sigma(k\alpha)) = \tau[k\sigma(\alpha)] = k\tau\sigma(\alpha).$$

因此 $\tau\sigma$ 是 V 到 V'' 的一个同构映射.　x

习　题　8.3

1. 证明:实数域 R 作为它自身上的线性空间与习题8.1第1题(2)中的线性空间 R^+ 同构.

2. 证明:域 F 上的线性空间 $M_{s\times n}(F)$ 与 F^{sn} 同构,并且写出一个同构映射.

3. 证明:域 F 上次数 $< n$ 的一元多项式组成的线性空间 $F[x]_n$ 与 F^n 同构,并且写出一个同构映射.

4. 设域 F 上线性空间 V 中的向量组 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 线性无关,试问:向量组 $\alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \alpha_3 + \alpha_4, \alpha_4 + \alpha_1$ 是否线性无关?令

$$W = \langle \alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \alpha_3 + \alpha_4, \alpha_4 + \alpha_1 \rangle,$$

求 W 的一个基和维数.

5. 令

$$L = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \middle| a, b \in R \right\},$$

(1) 证明 L 是实线性空间 $M_2(R)$ 的子空间,求 L 的一个基和维数;

(2) 证明复数域 C 作为实数域 R 上的线性空间与 L 同构,并且写出一个同构映射.

*6. 令

$$H = \left\{ \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \middle| \alpha, \beta \in C \right\},$$

(1) 证明 H 对于矩阵的加法,以及实数与矩阵的数量乘法(注意实数可看成复数)构成实数域上的一个线性空间;

(2) 求 H 的一个基和维数;

(3) 证明 H 与 R^4 同构,并且写出一个同构映射.

阅读材料四:有限域的元素个数

设 F 是一个有限域,它的单位元为 e.如果 F 的特征为 0,则 $e, 2e, \cdots, ne, \cdots$ 是 F 中两两不同的元素,从而 F 有无穷多个元素,矛盾.因此 F 的特征为一个素数 p.令

$$F_p = \{0, e, 2e, \cdots, (p-1)e\}.$$

易证 F_p 对于 F 的减法、乘法封闭,从而 F_p 是 F 的一个子环,显然 e 是 F_p 的一个单位元,且 F_p 为交换环,任取 F_p 的一个非零元 ie,由于 $(i, p) = 1$,因此存在 $u, v \in Z$,使得 $ui + vp = 1$.从而

$$e = 1e = uie + vpe = u(ie) = (ue)(ie).$$

设 $u = lp + r, \quad 0 \leq r < p$,则

$$ue = (lp + r)e = lpe + re = re \in F_p.$$

因此 ie 在 F_p 中有逆元 re,从而 F_p 是一个域.

有限域 F 可看成是域 F_p 上的线性空间,其中加法是 F 的加法,纯量乘法是 F_p 中元素与 F 中元素做 F 的乘法.由于 F 是域,因此它满足线性空间的定义中的 8 条运算法则,从而 F 是域 F_p 上的线性空间,由于 F 只含有限多个元素,因此 F 一定是有限维的,设其维数为 n,则 $F \cong F_p^n$.于是 F 到 F_p^n 有一个双射,从而 $|F| = |F_p^n|$.由于 $F_p^n = \{(a_1, a_2, \cdots, a_n) \mid a_i \in F_p, i = 1, 2, \cdots, n\}$,因此 $|F_p^n| = p \cdot p \cdot \cdots \cdot p = p^n$.从而 $|F| = p^n$,这样我们证明了

定理 1 有限域的元素个数是它的特征 p 的方幂,其中 p 是素数. □

§4 商 空 间

几何空间可以看成是由原点 O 为起点的所有向量组成的 3 维实线性空间 V,过原点的一个平面 W 是 V 的一个 2 维子空间,与 W 平行的每一个平面 π 不是 V 的子空间,因为 π 对于向量的加法和数量乘法都不封闭.自然要问:π 具

有什么样的结构？π 与 W 的关系如何？

在 π 上取定一个向量 γ_0，如图 8 - 3 所示，π 上每一个向量 γ 可以唯一地表示成 γ_0 与 W 中一个向量 η 之和：

$$\gamma = \gamma_0 + \eta.$$

反之，任取 $\eta \in W$，有 $\gamma_0 + \eta \in \pi$，因此

$$\pi = \{\gamma_0 + \eta \mid \eta \in W\}. \qquad (1)$$

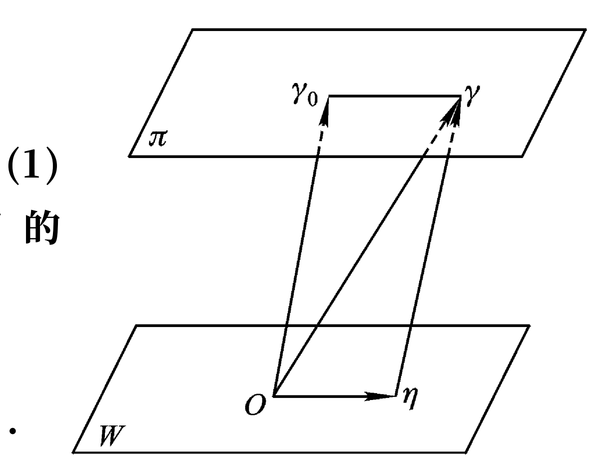


图 8 - 3

把(1)式右边的集合记作 $\gamma_0 + W$，称它是 W 的一个陪集，γ_0 称为陪集代表。显然

$$\gamma \in \gamma_0 + W \qquad \gamma = \gamma_0 + \eta$$
$$\gamma - \gamma_0 = \eta \in W.$$

由此看出，如果在 V 上规定一个二元关系～如下：

$$\gamma \sim \gamma_0 \qquad \gamma - \gamma_0 \in W, \qquad (2)$$

那么容易验证～具有反身性、对称性和传递性（因为 W 是 V 的一个子空间），从而～是等价关系，于是 γ_0 所属的等价类 $\bar{\gamma}_0$ 为

$$\begin{aligned}\bar{\gamma}_0 &= \{\gamma \in V \mid \gamma \sim \gamma_0\} = \{\gamma \in V \mid \gamma - \gamma_0 \in W\} \\ &= \{\gamma \in V \mid \gamma = \gamma_0 + \eta, \eta \in W\} \\ &= \{\gamma_0 + \eta \mid \eta \in W\} = \gamma_0 + W.\end{aligned}$$

这表明陪集 $\gamma_0 + W$ 是等价类 $\bar{\gamma}_0$.

综上所述，在几何空间 V 中，与 W 平行的每一个平面 π 是 W 的一个陪集，也就是在(2)式规定的等价关系下的一个等价类.

一般地，设 V 是域 F 上的一个线性空间，W 是 V 的一个子空间.在 V 上定义一个二元关系～如下：

$$\alpha \sim \beta \qquad \alpha - \beta \in W \qquad (3)$$

则～是一个等价关系，理由如下：

任取 $\alpha \in V$，因为 $\alpha - \alpha = 0 \in W$，所以 $\alpha \sim \alpha$.

假设 $\alpha \sim \beta$，则 $\alpha - \beta \in W$，从而

$$\beta - \alpha = -(\alpha - \beta) \in W,$$

因此 $\beta \sim \alpha$.

类似地，还可以证明～具有传递性（利用 W 对加法封闭）.

α 的等价类 $\bar{\alpha}$ 为

$$\begin{aligned}\bar{\alpha} &= \{\beta \in V \mid \beta \sim \alpha\} = \{\beta \in V \mid \beta - \alpha \in W\} \\ &= \{\beta \in V \mid \beta = \alpha + \gamma, \gamma \in W\}\end{aligned}$$

$$= \{\alpha + \gamma \mid \gamma \in W\}, \tag{4}$$

把(4)式最后一个集合记作 $\alpha + W$,称它为 W 的一个陪集,α 称为陪集代表.于是 α 的等价类 $\bar{\alpha}$ 就是以 α 为代表的 W 的一个陪集 $\alpha + W$.从而

$$\beta \in \alpha + W \qquad \beta \sim \alpha \qquad \beta - \alpha \in W. \tag{5}$$

根据本套教材上册第 5 章§1 的事实 4 得出

$$\alpha + W = \delta + W \qquad \alpha - \delta \in W. \tag{6}$$

由(6)式看出,一个陪集 $\alpha + W$ 的陪集代表不唯一;如果 $\alpha - \delta \in W$,那么 δ 也可以作为这个陪集的代表.

根据第 5 章§1 的定义 4, V 的所有等价类(即 W 的所有陪集)组成的集合称为 V 对于关系~的商集,记作 $V/\sim$.由于关系~是利用子空间 W 确定的,因此把 $V/\sim$ 称为 V 对于子空间 W 的商集,记成 V/W,即

$$V/W = \{\alpha + W \mid \alpha \in V\}. \tag{7}$$

线性空间 V 是具有加法与纯量乘法两种运算的集合,因此我们有理由期望 V 对子空间 W 的商集 V/W 也可以定义加法与纯量乘法两种运算.但是由于 V/W 不是 V 的子集,因此不能把 V 的两种运算直接搬到 V/W 中来.

定理 1 设 W 是域 F 上线性空间 V 的一个子空间.在商集 V/W 中规定

$$(\alpha + W) + (\beta + W) \overset{\text{def}}{=} (\alpha + \beta) + W, \tag{8}$$

$$k(\alpha + W) \overset{\text{def}}{=} k\alpha + W, \quad k \in F, \tag{9}$$

则(8)式和(9)式分别定义了 V/W 中的加法运算与纯量乘法运算,并且 V/W 对于这两种运算成为域 F 上一个线性空间,它的零元素是 W(即 $0 + W$),线性空间 V/W 称为 V 对 W 的商空间.

证明 (8)式和(9)式都用到陪集代表,我们需要说明它们的结果不依赖于陪集代表的选择.设

$$\alpha_1 + W = \alpha + W, \qquad \beta_1 + W = \beta + W.$$

则 $\alpha_1 - \alpha \in W$, $\beta_1 - \beta \in W$.从而

$$(\alpha_1 + \beta_1) - (\alpha + \beta) = (\alpha_1 - \alpha) + (\beta_1 - \beta) \in W,$$

$$k\alpha_1 - k\alpha = k(\alpha_1 - \alpha) \in W,$$

因此

$$(\alpha_1 + \beta_1) + W = (\alpha + \beta) + W, \quad k\alpha_1 + W = k\alpha + W.$$

这证明了上述关于加法和纯量乘法运算的定义是合理的.

容易验证线性空间定义中的 8 条运算法则在 V/W 中都成立.例如

$$(\alpha + W) + (0 + W) = (\alpha + 0) + W = \alpha + W, \quad \forall\, \alpha + W \in V/W,$$

因此 $0 + W$(即 W)是 V/W 的零元素.

综上所述, V/W 成为域 F 上一个线性空间. □

例如,本节开头讲的几何空间 V 的例子中,与 W 平行的所有平面以及 W

组成的集合就是 V 对于 W 的商空间 V/W. 显然,与 W 平行的平面都可以由 W 沿 γ_0 或 $-\gamma_0$ 的方向平移得到(参看图 8-3),因此与 W 平行的每一个平面(即 V/W 中的每一个陪集)可写成 $k\gamma_0+W$ 的形式,其中 $k\in R$. 由于

$$k\gamma_0+W=k(\gamma_0+W),$$

因此 V/W 中每一个陪集 $k\gamma_0+W$ 可以由陪集 γ_0+W 线性表出,且 γ_0+W 线性无关,于是 γ_0+W 就是商空间 V/W 的一个基,从而 $\dim(V/W)=1$. 显然 $\dim V=3$, $\dim W=2$,由此看出, $\dim(V/W)=\dim V-\dim W$. 一般地,我们有下述结论:

定理 2　如果 W 是域 F 上有限维线性空间 V 的一个子空间,则

$$\dim(V/W)=\dim V-\dim W. \tag{10}$$

证明　在 W 中取一个基 $\alpha_1,\cdots,\alpha_s$,把它扩充成 V 的一个基

$$\alpha_1,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n.$$

任取 $\beta+W\in V/W$,设 $\beta=b_1\alpha_1+\cdots+b_n\alpha_n$,则

$$\begin{aligned}\beta+W&=(b_1\alpha_1+\cdots+b_n\alpha_n)+W\\&=(b_1\alpha_1+W)+\cdots+(b_s\alpha_s+W)\\&\quad+(b_{s+1}\alpha_{s+1}+W)+\cdots+(b_n\alpha_n+W)\\&=W+\cdots+W+b_{s+1}(\alpha_{s+1}+W)\\&\quad+\cdots+b_n(\alpha_n+W)\\&=b_{s+1}(\alpha_{s+1}+W)+\cdots+b_n(\alpha_n+W).\end{aligned} \tag{11}$$

这表明 V/W 中任一向量可表成 $\alpha_{s+1}+W,\cdots,\alpha_n+W$ 的线性组合. 若能证 $\alpha_{s+1}+W,\cdots,\alpha_n+W$ 线性无关,则它就是 V/W 的一个基,从而

$$\dim(V/W)=n-s=\dim V-\dim W.$$

现在假设

$$k_1(\alpha_{s+1}+W)+\cdots+k_{n-s}(\alpha_n+W)=W, \tag{12}$$

则　$(k_1\alpha_{s+1}+\cdots+k_{n-s}\alpha_n)+W=W,$

从而　$k_1\alpha_{s+1}+\cdots+k_{n-s}\alpha_n\in W,$

于是　$k_1\alpha_{s+1}+\cdots+k_{n-s}\alpha_n=l_1\alpha_1+\cdots+l_s\alpha_s,$

即

$$l_1\alpha_1+\cdots+l_s\alpha_s-k_1\alpha_{s+1}-\cdots-k_{n-s}\alpha_n=0. \tag{13}$$

由于 $\alpha_1,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n$ 线性无关,所以

$$l_1=\cdots=l_s=k_1=\cdots=k_{n-s}=0$$

这证明了 $\alpha_{s+1}+W,\cdots,\alpha_n+W$ 是 V/W 中线性无关的向量组. x

数学中许多重要问题会遇到线性空间 V 及其子空间 W 都是无限维,而商空间 V/W 是有限维. 在这种情形,定理 2 不适用, $\dim(V/W)$ 的计算常常变成

非平凡的问题.

定义 1　设 W 是域 F 上线性空间 V 的一个子空间,如果 V 对 W 的商空间 V/W 是有限维的,则 $\dim(V/W)$ 称为子空间 W 在 V 中的余维数(codimension),记作 $\mathrm{codim}_V W$ 或 $\mathrm{codim}\, W$.

设 W 是域 F 上线性空间 V 的一个子空间,则 V 到商空间 V/W 有一个很自然的映射 $\pi: \alpha \longmapsto \alpha + W$,称它为标准映射或典范映射(canonical mapping). 显然它是满射.

习　题　8.4

1. 设 U, W 都是域 F 上线性空间 V 的子空间,证明:如果 $V = U \oplus W$,那么
$$V/U \cong W.$$

2. 设 V 是几何空间, U 是过原点 O 的一条直线,商空间 V/U 由哪些元素组成?求 V/U 的一个基和维数.

3. 设 V 是域 F 上的一个 n 维线性空间, $n \geq 3$, U 是 V 的一个 2 维子空间,用 Ω_1 表示 V 中包含 U 的所有 $n-1$ 维子空间组成的集合,用 Ω_2 表示商空间 V/U 的所有 $n-3$ 维子空间组成的集合.令
$$\begin{aligned}\sigma: \Omega_1 &\to \Omega_2\\ W &\longmapsto W/U,\end{aligned}$$
证明: σ 是双射.

4. 设 U, W 都是域 F 上线性空间 V 的子空间,证明
$$U + W/W \cong U/U \cap W.$$

阅读材料五:线性码

在信息时代里,大量的信息要及时传递.例如,人造地球卫星拍摄的照片要及时发送回地面接收站,这需要利用无线电波.而工程上容易实现的是把无线电信号区分成两种状态,让一种状态对应于 0,另一种状态对应于 1.于是首先要把待发送的消息编成由 0 和 1 组成的字符串,然后利用无线电波发送.在传送过程中,受自然界的电磁源以及其他无线电系统发射的信号的干扰,有可能 0 错成 1,1 错成 0.这样接收到的字符串就不一定是原来发送的字符串.试问:能否检查出有无差错?如果发现有差错,能否纠正差错?运用线性空间的理论可以提供一种检错和纠错的方法.下面我们通过简单的例子来阐述这种方法.

把待发送的消息编成由 0 和 1 组成的 4 位字符串,并且把这种 4 位字符串

看成是二元域 Z_2 上的 4 维向量空间 Z_2^4 里的一个向量(a_1, a_2, a_3, a_4).如何察觉在传送过程中有无发生差错？从日常生活的例子可受到启发.例如,写一封英文信,如果一个单词的字母较多,那么容易察觉出拼写差错,譬如,communication(通信)如果写成"conmunication",那么容易察觉第三个字母"n"是错的.这表明一个单词如果有冗余度,那么就有可能察觉出拼写差错.由此受到启发,在每一个 4 维向量(a_1, a_2, a_3, a_4)的右边添上 3 个分量,成为 Z_2 上的 7 维向量$(a_1, a_2, a_3, a_4, c_1, c_2, c_3)$,其中

$$\begin{aligned} c_1 &= a_1 + a_2 + a_3, \\ c_2 &= a_1 + a_2 + a_4, \\ c_3 &= a_1 + a_3 + a_4. \end{aligned} \tag{1}$$

这样就给出了二元域 Z_2 上的向量空间 Z_2^4 到 Z_2^7 的一个映射 σ:

$$\begin{aligned} \sigma:\quad & Z_2^4 \longrightarrow Z_2^7 \\ & (a_1, a_2, a_3, a_4) \longmapsto (a_1, a_2, a_3, a_4, c_1, c_2, c_3), \end{aligned} \tag{2}$$

其中 c_1, c_2, c_3 如(1)式所示,显然 σ 是单射.称 σ 是一个编码;σ 的象 $\operatorname{Im}\sigma$ 称为一个码,通常记作 C;码 C 里的每一个元素称为一个码字;而 σ 的陪域 Z_2^7 的每一个元素称为一个字,码字$(a_1, a_2, a_3, a_4, c_1, c_2, c_3)$的前 4 个分量称为信息位,后 3 个分量称为校验位.现在我们来研究码 C 具有什么样的结构.用矩阵可以把(1)式写成

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}. \tag{3}$$

把(3)式右边的 3×4 矩阵记作 A,则(3)式可写成

$$A \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} - I_3 \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = 0. \tag{4}$$

(4)式等价于

$$(A \quad -I_3) \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} = 0. \tag{5}$$

把(5)式左边的分块矩阵$(A \quad -I_3)$记成 H,即

$$H = (A \quad -I_3). \tag{6}$$

则从(5)、(6)式得,对于 $\alpha \in Z_2^7$(元素写成行向量),有

$$\alpha \in C \qquad H\alpha' = 0. \tag{7}$$

这表明 α 是码字当且仅当 α'是齐次线性方程组 $HX=0$ 的一个解.由于 $HX=0$ 的解集 W 是 Z_2^7(元素写成列向量)的一个线性子空间,因此码 C是 Z_2^7(元素写成行向量)的一个线性子空间.由于 rank$(H)=3$,因此 $\dim W = 7-3=4$.从而

$$\dim C = 4. \tag{8}$$

我们称码 C 是(7,4)线性码,其中 7 是编码 σ 的陪域 Z_2^7 的维数,4 是码 C 的维数.(6)式给出的 3×7 矩阵 H 称为码 C 的校验矩阵.

设发送一个码字 α,接收的是字 γ.计算 $H\gamma'$,如果 $H\gamma' \neq 0$,则 γ 不是码字.从而察觉出传送过程中发生了差错.这时能否纠正差错,从 γ 恢复成发送的码字 α?由于传送码字的途径(称为信道)总应该是:出错少的可能性较大,出错多的可能性较小,因此应当在码 C 中寻找一个码字,它与 γ 的对应分量不同的位置最少.为此我们引出一个概念:

定义 1　设 $\alpha, \beta \in Z_2^n$,α 与 β 对应分量不同的位置的个数称为 α 与 β 的 Hamming 距离,记作 $d(\alpha, \beta)$.

显然 $d(\alpha, \beta)$等于向量 $\alpha-\beta$ 的非零分量的个数,为此又引出一个概念:

定义 2　设 $\alpha \in Z_2^n$,α 的非零分量的个数称为 α 的 Hamming 重量,记作 $W(\alpha)$.

从上面所说知道,

$$d(\alpha, \beta) = W(\alpha - \beta). \tag{9}$$

如果收到的字 γ 不是码字,那么我们去求 C 中每一个码字与 γ 的 Hamming 距离,从中找出与 γ 的 Hamming 距离最短的码字 β,把 γ 译成这个码字 β.在编码 σ 满足一定条件下,β 很可能就是原来发送的码字 α.这种译码想法称为极大似然译码原理.为了减少上述计算量,我们来仔细分析,收到的字 γ 与发送的码字 α 之间的关系.令

$$e = \gamma - \alpha, \tag{10}$$

称 e 是差错向量.我们有

$$He' = H(\gamma - \alpha)' = H\gamma' - H\alpha' = H\gamma', \tag{11}$$

称 $H\gamma'$是 γ 的校验子.从(11)式看出,差错向量 e 与收到的字 γ 有相同的校验子.

一般地,　　α 与 β 有相同的校验子

$$H\alpha' = H\beta'$$

$$H(\alpha-\beta)'=0$$
$$\alpha-\beta\in C$$
$$\alpha+C=\beta+C. \tag{12}$$

这表明 α 与 β 有相同的校验子当且仅当 α 与 β 属于码 C 的同一个陪集,由此得出,差错向量 e 属于陪集 $\gamma+C$.由极大似然译码原理得,陪集 $\gamma+C$ 中重量最小的向量最有可能是差错向量 e.于是我们把 γ 就译成码字 $\gamma-e$.陪集中重量最小的向量称为陪集头.

在实际译码中,把码 C 的所有码字排在第一行,其余每个陪集的向量排成一行.求出每一个陪集的陪集头,写在该行的最左边;求出该陪集头的校验子,写在该行的最右边,得到一张译码表.收到一个字 γ 后,计算它的校验子 $H\gamma'$,从译码表的最右边一列查出该校验子,从这个校验子所在的行查出字 γ,从 γ 所在的列找出第一行里的码字,则把 γ 就译成这个码字.

为简单起见,我们举一个线性码 C 作为例子.设码 C 的校验矩阵 H 为

$$H=\begin{pmatrix}1&1&1&0\\0&1&0&1\end{pmatrix}.$$

容易看出,$\mathrm{rank}(H)=2$,因此齐次线性方程组 $HX=0$ 的解空间 W 的维数为 $4-2=2$.从而 $\dim C=2$.于是 C 是(4,2)线性码,由于商空间 Z_2^4/C 的维数为

$$\dim(Z_2^4/C)=\dim Z_2^4-\dim C=4-2=2,$$

因此商空间 Z_2^4/C 的元素个数为 $2^2=4$,即 C 的陪集共有 4 个.现在可以列出码 C 的译码表如下(把 (a_1,a_2,a_3,a_4) 简写成 $a_1a_2a_3a_4$):

陪集头	陪集里的其他元素	校验子
0000	1010　0111　1101	$\begin{pmatrix}0\\0\end{pmatrix}$
1000	0010　1111　0101	$\begin{pmatrix}1\\0\end{pmatrix}$
0100	1110　0011　1001	$\begin{pmatrix}1\\1\end{pmatrix}$
0001	1011　0110　1100	$\begin{pmatrix}0\\1\end{pmatrix}$

例如,设收到的字 $\gamma=1110$,计算校验子 $H\gamma'=\begin{pmatrix}1\\1\end{pmatrix}$.从译码表的最右边一列查出 $\begin{pmatrix}1\\1\end{pmatrix}$,在该行里找到 1110,它所在列的第一行中的码字为 1010,因此把 γ 就译成码字 1010.

第 9 章　线 性 映 射

我们在上一章研究了域 F 上线性空间的结构.在许多数学分支和实际问题中都会遇到线性空间之间的映射,并且这种映射保持加法和纯量乘法两种运算,称它为线性映射.线性代数就是研究线性空间和线性映射的理论.这一章我们来研究线性映射的理论.

§1　线性映射及其运算

几何空间 V 中,把每个向量 $\alpha=(x, y, z)$ 对应到 $\alpha_1=(x, y, 0)$ 的对应法则是 V 到 V 的一个映射,称它为 V 在 xOy 平面(记作 W)上的正投影,记作 P_W,即

$$P_W(\alpha)=\alpha_1,$$

如图 9 - 1 所示.显然 P_W 保持加法和数量乘法两种运算.

闭区间 $[a, b]$ 上所有连续函数组成的集合 $C[a, b]$ 对于函数的加法,以及数与函数的数量乘法成为实数域 R 上的线性空间.函数的定积分是域 R 上的线性空间 $C[a, b]$ 到 R 的一个映射,它具有下述性质:

$$\int_a^b [f(x)+g(x)]dx=\int_a^b f(x)dx+\int_a^b g(x)dx,$$

$$\int_a^b kf(x)dx=k\int_a^b f(x)dx,$$

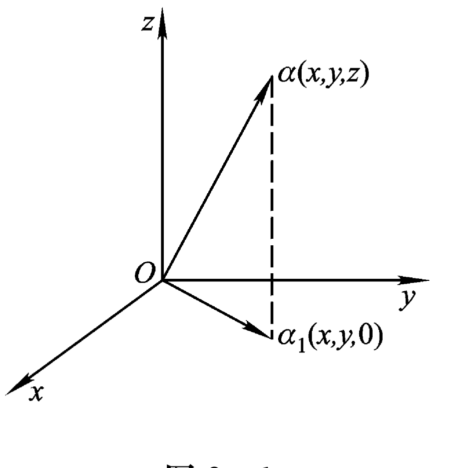


图 9 - 1

即,函数的定积分保持加法和数量乘法两种运算.

定义 1　设 V 与 V' 是域 F 上的两个线性空间,V 到 V' 的一个映射 A 如果保持加法运算和纯量乘法运算,即

$$A(\alpha+\beta)=A(\alpha)+A(\beta),\quad \forall\,\alpha, \beta\in V \tag{1}$$

$$A(k\alpha)=kA(\alpha),\quad \forall\,\alpha\in V, k\in F \tag{2}$$

则称 A 是 V 到 V' 的一个线性映射.

线性空间 V 到自身的线性映射通常称为 V 上的线性变换.

域 F 上的线性空间 V 到 F 的线性映射称为 V 上的线性函数.

例如,几何空间 V 在 xOy 平面上的正投影 P_W 是 V 上的一个线性变换.函数的定积分是 $C[a,b]$上的一个线性函数.

例 1　线性空间 V 到 V'的零映射(即 $A(\alpha)=0$, $\forall\alpha\in V$)是线性映射,记作 0.

例 2　线性空间 V 上的恒等变换 1_V 是 V 上的一个线性变换,也可记成 I.

例 3　给定 $k\in F$, F 上线性空间 V 到自身的一个映射 $k(\alpha)=k\alpha$,称为 V 上由 k 决定的数乘变换.当 $k=0$ 时,便得到零变换;当 $k=1$ 时,便得到恒等变换.

例 4　设 A 是域 F 上的一个 $s\times n$ 矩阵,令

$$A(\alpha)\overset{\text{def}}{=}A\alpha,\ \forall\alpha\in F^n,$$

则 A 是 F^n 到 F^s 的一个线性映射.

例 5　用 $C^{(1)}(a,b)$表示区间(a,b)上所有一次可微函数组成的集合,容易验证它是实数域 R 上的线性空间 $R^{(a,b)}$ 的一个子空间.用 D 表示求导数,即

$$D(f(x))=f'(x),\ \forall f(x)\in C^{(1)}(a,b),$$

则 D 是 $C^{(1)}(a,b)$到 $R^{(a,b)}$ 的一个线性映射.

例 6　σ是线性空间 V 到 V'的一个同构映射当且仅当σ是 V 到 V'的一个可逆线性映射.

证明　由于同构映射比线性映射多一个条件:双射,而一个映射是双射当且仅当它是可逆映射,因此命题成立.　x

由于线性映射只比同构映射少了双射这一条件,因此同构映射的性质中,只要它的证明没有用到单射和满射的条件,那么对于线性映射也成立.于是如果 A 是数域 K 上线性空间 V 到 V'的线性映射,则 A 有下述性质:

1° $A(0)=0'$,其中 $0'$是 V'的零向量;

2° $A(-\alpha)=-A(\alpha)$, $\forall\alpha\in V$;

3° $A(k_1\alpha_1+k_2\alpha_2+\cdots+k_s\alpha_s)=k_1A(\alpha_1)+k_2A(\alpha_2)+\cdots+k_sA(\alpha_s)$;

4°如果 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是 V 的一个线性相关的向量组,则 $A(\alpha_1)$, $A(\alpha_2)$,$\cdots$, $A(\alpha_s)$是 V'的一个线性相关的向量组;但是反之不成立.即线性映射可能把线性无关的向量组变成线性相关的向量组;

5°如果 V 是有限维的,且 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 的一个基,则对于 V 中任一向量 $\alpha=a_1\alpha_1+a_2\alpha_2+\cdots+a_n\alpha_n$,有

$$A(\alpha)=a_1A(\alpha_1)+a_2A(\alpha_2)+\cdots+a_nA(\alpha_n).\tag{3}$$

这表明,只要知道了 V 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 在 A 下的象,那么 V 中任一向量在 A 下的象就都确定了.即 n 维线性空间 V 到 V'的线性映射完全被它在 V 的一个基上的作用所决定.

给了域 F 上任意两个线性空间 V 和 V',是否存在 V 到 V' 的一个线性映射? 如果 V 是有限维的,那么回答是肯定的.即我们有下述结论:

定理 1 设 V 和 V' 都是域 F 上的线性空间, V 的维数是 n, V 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$, V' 中任意取定 n 个向量 $\gamma_1,\gamma_2,\cdots,\gamma_n$(它们中可以有相同的),令

$$\mathbf{A}: \quad V \to V'$$

$$\alpha=\sum_{i=1}^{n}a_i\alpha_i \longmapsto \sum_{i=1}^{n}a_i\gamma_i, \tag{4}$$

则 $\mathbf{A}$ 是 V 到 V' 的一个线性映射,且 $\mathbf{A}(\alpha_i)=\gamma_i$, $i=1,2,\cdots,n$.

证明 由于 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是 V 的一个基,因此 α 表示成 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的线性组合的方式唯一,从而 $\mathbf{A}$ 是 V 到 V' 的一个映射.在 V 中任取两个向量

$$\alpha=\sum_{i=1}^{n}a_i\alpha_i, \qquad \beta=\sum_{i=1}^{n}b_i\alpha_i,$$

则

$$\begin{aligned}\mathbf{A}(\alpha+\beta)&=\mathbf{A}\left(\sum_{i=1}^{n}(a_i+b_i)\alpha_i\right)=\sum_{i=1}^{n}(a_i+b_i)\gamma_i\\&=\sum_{i=1}^{n}a_i\gamma_i+\sum_{i=1}^{n}b_i\gamma_i=\mathbf{A}(\alpha)+\mathbf{A}(\beta),\\ \mathbf{A}(k\alpha)&=\mathbf{A}\left(\sum_{i=1}^{n}(ka_i)\alpha_i\right)=\sum_{i=1}^{n}(ka_i)\gamma_i\\&=k\sum_{i=1}^{n}a_i\gamma_i=k\mathbf{A}(\alpha), \qquad \forall k\in F.\end{aligned}$$

因此 $\mathbf{A}$ 是 V 到 V' 的一个线性映射.显然有

$$\mathbf{A}(\alpha_i)=\mathbf{A}(0\alpha_1+\cdots+0\alpha_{i-1}+1\alpha_i+0\alpha_{i+1}+\cdots+0\alpha_n)=\gamma_i,$$

其中 $i=1,2,\cdots,n$. $\blacksquare$

由于 V 到 V' 的线性映射完全被它在 V 的一个基上的作用所决定,因此定理 1 中满足 $\mathbf{A}(\alpha_i)=\gamma_i$ $(i=1,2,\cdots,n)$的线性映射是唯一的.

定理 2 设 V 是域 F 上的一个线性空间, U, W 是 V 的两个子空间,且

$$V=U\oplus W. \tag{5}$$

任取 $\alpha\in V$,设 $\alpha=\alpha_1+\alpha_2$,其中 $\alpha_1\in U$, $\alpha_2\in W$.令

$$\mathbf{P}_U: \quad V \to V$$

$$\alpha=\alpha_1+\alpha_2 \longmapsto \alpha_1, \tag{6}$$

则 $\mathbf{P}_U$ 是 V 上的一个线性变换,称 $\mathbf{P}_U$ 是平行于 W 在 U 上的投影,它满足

$$\mathbf{P}_U(\alpha)=\begin{cases}\alpha, & \text{当 } \alpha\in U,\\ 0, & \text{当 } \alpha\in W;\end{cases} \tag{7}$$

满足(7)式的 V 上的线性变换是唯一的.

证明　由于 $V = U \oplus W$，因此 α 表示成 U 的一个向量与 W 的一个向量之和的方式唯一．从而 P_U 是 V 到 V 的一个映射．任取 V 中两个向量

$$\alpha = \alpha_1 + \alpha_2, \qquad \beta = \beta_1 + \beta_2,$$

其中 $\alpha_1, \beta_1 \in U, \alpha_2, \beta_2 \in W$．则 $\alpha_1 + \beta_1 \in U, \alpha_2 + \beta_2 \in W$．从而

$$P_U(\alpha + \beta) = P_U[(\alpha_1 + \beta_1) + (\alpha_2 + \beta_2)] = \alpha_1 + \beta_1 = P_U(\alpha) + P_U(\beta),$$

$$P_U(k\alpha) = P_U(k\alpha_1 + k\beta_1) = k\alpha_1 = kP_U(\alpha), \quad \forall k \in F,$$

因此 P_U 是 V 上的一个线性变换．

如果 $\alpha \in U$，则 $\alpha = \alpha + 0$，从而 $P_U(\alpha) = \alpha$．

如果 $\alpha \in W$，则 $\alpha = 0 + \alpha$，从而 $P_U(\alpha) = 0$．

设 V 上的线性变换 A 也满足(7)式，任取 $\alpha \in V$，设 $\alpha = \alpha_1 + \alpha_2$，其中 $\alpha_1 \in U$，$\alpha_2 \in W$，则

$$A(\alpha) = A(\alpha_1 + \alpha_2) = A(\alpha_1) + A(\alpha_2) = \alpha_1 + 0 = \alpha_1 = P_U(\alpha),$$

因此，$A = P_U$．　□

类似地，定义 $P_W(\alpha) = \alpha_2$，则 P_W 也是 V 上的一个线性变换，称它为平行于 U 在 W 上的投影．

定理 2 告诉我们，如果线性空间 V 能分解成两个子空间的直和：$V = U \oplus W$，那么 V 在子空间 U(或 W)上的投影 P_U(或 P_W)就是 V 上的一个线性变换，且 P_U 满足(7)式．投影是非常重要的一类线性变换．

V 在子空间 U(或 W)上的投影 P_U(或 P_W)是 V 到 V 的一个映射，根据映射的乘法，有 $P_U^2 (= P_U P_U)$，$P_U P_W$，$P_W P_U$，P_W^2．任取 $\alpha \in V$，设 $\alpha = \alpha_1 + \alpha_2$，$\alpha_1 \in U$，$\alpha_2 \in W$．则

$$P_U^2(\alpha) = P_U(P_U(\alpha)) = P_U(\alpha_1) = \alpha_1 = P_U(\alpha),$$

$$P_U P_W(\alpha) = P_U(P_W(\alpha)) = P_U(\alpha_2) = 0,$$

$$P_W P_U(\alpha) = P_W(P_U(\alpha)) = P_W(\alpha_1) = 0,$$

因此

$$P_U^2 = P_U, \qquad P_U P_W = P_W P_U = 0 \tag{8}$$

类似地有 $P_W^2 = P_W$．

V 上的线性变换 A 如果满足 $A^2 = A$，则称 A 是幂等变换．

V 上的两个线性变换 A, B 如果满足 $AB = BA = 0$，则称 A 与 B 是正交的．

从(8)式知道，如果 $V = U \oplus W$，则投影 P_U, P_W 都是幂等变换，而且 P_U 与 P_W 是正交的．

(8)式表明，我们可以通过线性变换的运算来刻画线性变换的性质．下面我们就来研究线性映射的运算．

今后我们用 $A\alpha$ 来记 $A(\alpha)$,使书写简洁一些.

命题 3 设 V, U, W 都是域 F 上的线性空间, A 是 V 到 U 的一个线性映射, B 是 U 到 W 的一个线性映射,则 BA 是 V 到 W 的一个线性映射.

证明 显然 BA 是 V 到 W 的一个映射.任取 $\alpha, \beta \in V$, $k \in F$,有

$$
\begin{aligned}
(BA)(\alpha+\beta) &= B(A(\alpha+\beta)) = B(A\alpha + A\beta) \\
&= B(A\alpha) + B(A\beta) = (BA)\alpha + (BA)\beta, \\
(BA)(k\alpha) &= B(A(k\alpha)) = B(kA\alpha) \\
&= k(B(A\alpha)) = k((BA)\alpha),
\end{aligned}
$$

因此 BA 是 V 到 V' 的一个线性映射. x

由于映射的乘法适合结合律,不适合交换律,因此线性映射的乘法也适合结合律,不适合交换律.

命题 4 设 A 是线性空间 V 到 V' 的一个线性映射,如果 A 可逆,则 A^{-1} 是 V' 到 V 的一个线性映射.

证 A 是 V 到 V' 的可逆线性映射

A 是 V 到 V' 的同构映射

A^{-1} 是 V' 到 V 的同构映射

A^{-1} 是 V' 到 V 的可逆线性映射. x

由于域 F 上两个有限维线性空间 V 与 V' 同构的充分必要条件是它们的维数相同,因此 V 到 V' 的可逆线性映射存在的充分必要条件是 $\dim V = \dim V'$.

线性映射还有加法运算与纯量乘法运算.

命题 5 设 A, B 都是域 F 上线性空间 V 到 V' 的线性映射, $k \in F$,令

$$(A+B)\alpha \overset{\text{def}}{=} A\alpha + B\alpha, \quad \forall \alpha \in V; \tag{9}$$

$$(kA)\alpha \overset{\text{def}}{=} k(A\alpha), \quad \forall \alpha \in V, \tag{10}$$

则 $A+B$, kA 都是 V 到 V' 的线性映射,称 $A+B$ 是 A 与 B 的和,称 kA 是 k 与 A 的纯量乘积.

证明 显然由(9)式定义的 $A+B$ 是 V 到 V' 的一个映射.对于任意 $\alpha, \beta \in V$, $l \in F$,有

$$
\begin{aligned}
(A+B)(\alpha+\beta) &= A(\alpha+\beta) + B(\alpha+\beta) \\
&= A\alpha + A\beta + B\alpha + B\beta \\
&= (A+B)\alpha + (A+B)\beta, \\
(A+B)(l\alpha) &= A(l\alpha) + B(l\alpha) = lA\alpha + lB\alpha \\
&= l(A\alpha + B\alpha) = l(A+B)\alpha,
\end{aligned}
$$

因此 $A+B$ 是 V 到 V' 的线性映射.

同理可证由(10)式定义的 kA 是 V 到 V' 的线性映射. x

容易验证,线性映射的加法与纯量乘法满足线性空间定义中的 8 条运算法则,因此域 F 上线性空间 V 到 V'的所有线性映射组成的集合成为域 F 上的一个线性空间,记作 $Hom(V, V')$.

特别地,域 F 上线性空间 V 上的所有线性变换组成的集合成为域 F 上的一个线性空间,记作 $Hom(V, V)$.在 $Hom(V, V)$中还有乘法运算.我们已经知道,线性变换的乘法满足结合律;恒等变换 I 具有性质: $IA = AI = A$, $\forall A \in Hom(V, V)$.容易验证,线性变换的乘法对于加法还满足左、右分配律:

$$A(B + C) = AB + AC,$$

$$(B + C)A = BA + CA.$$

因此 $Hom(V, V)$对于加法和乘法成为一个有单位元的环.容易验证,线性变换的乘法与纯量乘法满足:

$$k(AB) = (kA)B = A(kB), \tag{11}$$

其中 $k \in F$, $A, B \in Hom(V, V)$.

定义 2　一个非空集合 A 如果有加法、乘法运算,以及域 F 与 A 的纯量乘法运算,并且 A 对于加法和纯量乘法成为域 F 上的一个线性空间, A 对于加法和乘法成为一个有单位元的环, A 的乘法与纯量乘法满足:

$$k(\alpha\beta) = (k\alpha)\beta = \alpha(k\beta), \ \forall k \in F, \alpha, \beta \in A, \tag{12}$$

则称 A 是域 F 上的一个代数,并且把线性空间 A 的维数称为代数 A 的维数.

从上面的讨论知道,域 F 上线性空间 V 上的所有线性变换组成的集合 $Hom(V, V)$,对于线性变换的加法、乘法与纯量乘法,成为域 F 上的一个代数.

容易看出,域 F 上所有 n 级矩阵组成的集合 $M_n(F)$,对于矩阵的加法、乘法与纯量乘法,成为域 F 上的一个代数.

由于线性变换的乘法满足结合律,因此可以定义线性变换 A 的正整数指数幂:

$$A^m \stackrel{\text{def}}{=} \underbrace{A \cdot A \cdot \cdots \cdot A}_{m\text{个}}. \tag{13}$$

还可以定义 A 的零次幂:

$$A^0 \stackrel{\text{def}}{=} I. \tag{14}$$

容易验证:

$$A^m \cdot A^n = A^{m+n}, \quad (A^m)^n = A^{mn}, \quad m, n \in N. \tag{15}$$

当 A 可逆时,可以定义

$$A^{-m} \stackrel{\text{def}}{=} (A^{-1})^m, \ m \in N. \tag{16}$$

设 $f(x) = a_0 + a_1 x + \cdots + a_m x^m$ 是域 F 上的一元多项式, x 用 V 上的线性变换 A 代入,得

$$f(\mathbf{A}) = a_0\mathbf{I} + a_1\mathbf{A} + \cdots + a_m\mathbf{A}^m. \tag{17}$$

显然，$f(\mathbf{A})$仍是 V 上的一个线性变换，称 $f(\mathbf{A})$是 $\mathbf{A}$ 的一个多项式.容易验证，线性变换 $\mathbf{A}$ 的任意两个多项式 $f(\mathbf{A})$与 $g(\mathbf{A})$是可交换的，即

$$f(\mathbf{A})g(\mathbf{A}) = g(\mathbf{A})f(\mathbf{A}). \tag{18}$$

把 V 上线性变换 $\mathbf{A}$ 的所有多项式组成的集合记作 $F[\mathbf{A}]$.容易验证，$F[\mathbf{A}]$对于线性变换的减法、乘法都封闭，从而 $F[\mathbf{A}]$是环 $\mathrm{Hom}(V, V)$的一个子环，且 $F[\mathbf{A}]$是交换环，$\mathbf{I}\in F[\mathbf{A}]$. $F[\mathbf{A}]$中所有数乘变换组成的集合是 $F[\mathbf{A}]$的一个子环，并且域 F 与这个子环之间有一个双射 $\tau: k \longmapsto \mathbf{k}$，$\tau$ 保持加法与乘法运算，因此 $F[\mathbf{A}]$可看成是 F 的一个扩环.

利用线性变换的运算，我们可以研究线性变换的性质：例如，设 $V = U \oplus W$，任取 $\alpha\in V$，设

$$\alpha = \alpha_1 + \alpha_2, \quad \alpha_1 \in U, \alpha_2 \in V,$$

则
$$(\mathbf{P}_U + \mathbf{P}_W)\alpha = \mathbf{P}_U\alpha + \mathbf{P}_W\alpha = \alpha_1 + \alpha_2 = \alpha = \mathbf{I}\alpha,$$

因此
$$\mathbf{P}_U + \mathbf{P}_W = \mathbf{I}. \tag{19}$$

(18)和(19)式表明，如果 $V = U \oplus W$，则投影 $\mathbf{P}_U$ 与 $\mathbf{P}_W$ 是正交的幂等变换，且 $\mathbf{P}_U + \mathbf{P}_W$ 等于恒等变换 $\mathbf{I}$.这个结果是重要的.

习 题 9.1

1. 判断下面所定义的 K^3 上的变换，哪些是线性变换？

(1) $\mathbf{A}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_2 + x_3 \\ x_3^2 \end{pmatrix}$；　(2) $\mathbf{A}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 \\ x_2 + x_3 \\ 3x_1 - x_2 + x_3 \end{pmatrix}$.

2. 判断下面所定义的 $M_n(K)$上的变换，哪些是线性变换？

(1) 设 $A\in M_n(K)$，令

$$\mathbf{A}(X) = XA, \quad \forall X\in M_n(K);$$

(2) 设 $B, C\in M_n(K)$，令

$$\mathbf{A}(X) = BXC, \quad \forall X\in M_n(K).$$

3. 判断下面定义的 $K[x]$上的变换是不是线性变换？给定 $a\in K$，令

$$\mathbf{A}f(x) = f(x+a), \ \forall f(x)\in K[x].$$

4. 设 R^+ 是习题 8.1 第 1 题(2)小题的实线性空间，判别 R^+ 到 R 的下述映射是不是线性映射？设 $a > 0$ 且 $a\neq 1$，令

$$\log_a : R^+ \to R$$
$$x \longmapsto \log_a x.$$

*5. 设 V 是 $K[x,y]$ 中所有 m 次齐次多项式组成的集合，它对于多项式的加法，以及数与多项式的乘法成为数域 K 上的一个线性空间。给定数域 K 上的一个 2 级矩阵 $A=(a_{ij})$，定义 V 到自身的一个映射 $\mathcal{A}$ 如下：

$$(\mathcal{A}f)(x,y) \overset{\text{def}}{=} f(a_{11}x+a_{21}y, a_{12}x+a_{22}y),$$

判断 $\mathcal{A}$ 是不是 V 上的一个线性变换？

*6. 把 2^m 个元素的有限域 F_{2^m} 看成 F_2 上的线性空间，令

$$f(x)=x^2,$$

判断 f 是不是 F_{2^m} 上的一个线性变换？

7. 在 $K[x]$ 中，令

$$\mathcal{A}f(x)=xf(x), \quad \forall f(x)\in K[x].$$

(1) 证明 $\mathcal{A}$ 是 $K[x]$ 上的一个线性变换；

(2) 用 $\mathcal{D}$ 表示求导数，证明：

$$\mathcal{D}\mathcal{A}-\mathcal{A}\mathcal{D}=\mathcal{I}.$$

8. 设 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 是线性空间 V 的一个基，$\mathcal{A}$ 是 V 上的一个线性变换，证明：$\mathcal{A}$ 可逆当且仅当 $\mathcal{A}\alpha_1,\mathcal{A}\alpha_2,\cdots,\mathcal{A}\alpha_n$ 是 V 的一个基。

9. 设 $\mathcal{A}$ 是线性空间 V 上的一个线性变换，证明：如果

$$\mathcal{A}^{m-1}\alpha\neq 0, \quad \mathcal{A}^m\alpha=0 \quad (m>0),$$

则 $\alpha,\mathcal{A}\alpha,\mathcal{A}^2\alpha,\cdots,\mathcal{A}^{m-1}\alpha$ 线性无关。

10. 设 $\mathcal{A},\mathcal{B}$ 是 V 上的线性变换，如果 $\mathcal{A}\mathcal{B}-\mathcal{B}\mathcal{A}=\mathcal{I}$，证明：

$$\mathcal{A}^k\mathcal{B}-\mathcal{B}\mathcal{A}^k=k\mathcal{A}^{k-1}, \quad k\geq 1.$$

11. 设 V 是域 F 上的一个线性空间，$\operatorname{char} F\neq 2$。设 $\mathcal{A},\mathcal{B}$ 是 V 上的幂等变换，证明：

(1) $\mathcal{A}+\mathcal{B}$ 是幂等变换当且仅当 $\mathcal{A}\mathcal{B}=\mathcal{B}\mathcal{A}=0$；

(2) 如果 $\mathcal{A}\mathcal{B}=\mathcal{B}\mathcal{A}$，则 $\mathcal{A}+\mathcal{B}-\mathcal{A}\mathcal{B}$ 也是幂等变换。

§2 线性映射的核与象

几何空间 V 在 xOy 平面上的正投影 $\mathcal{P}_W$ 把 z 轴上的每一个向量 $\alpha=(0,0,z)$ 映到零向量 $0=(0,0,0)$，$\mathcal{P}_W$ 的象是 xOy 平面，参看图 9-1。由此我们引出下述概念：

定义 1 设 V 和 V' 都是域 F 上的线性空间，$\mathcal{A}$ 是 V 到 V' 的一个线性映射，我们把 V' 中零向量 0 在 $\mathcal{A}$ 下的原象集称为 $\mathcal{A}$ 的核，记作 $\operatorname{Ker}\mathcal{A}$，即

$$\operatorname{Ker}\mathcal{A} \overset{\text{def}}{=} \{\alpha\in V \mid \mathcal{A}\alpha=0\}; \tag{1}$$

我们把映射 $\mathcal{A}$ 的象（也叫做 $\mathcal{A}$ 的值域）记作 $\operatorname{Im}\mathcal{A}$ 或 $\mathcal{A}V$。

命题 1 设 $\mathcal{A}$ 是域 F 上线性空间 V 到 V' 的一个线性映射，则 $\operatorname{Ker}\mathcal{A}$ 是 V 的一个子空间，$\operatorname{Im}\mathcal{A}$ 是 V' 的一个子空间。

证明 由于 $\mathcal{A}(0)=0$,因此 $0\in \mathrm{Ker}\,\mathcal{A}$.任取 $\alpha,\beta\in \mathrm{Ker}\,\mathcal{A}$, $k\in F$,有

$$\mathcal{A}(\alpha+\beta)=\mathcal{A}\alpha+\mathcal{A}\beta=0+0=0,$$
$$\mathcal{A}(k\alpha)=k\mathcal{A}\alpha=k0=0,$$

因此 $\mathrm{Ker}\,\mathcal{A}$ 是 V 的一个子空间.

显然,$\mathrm{Im}\,\mathcal{A}$ 是 V'的一个非空子集.在 $\mathrm{Im}\,\mathcal{A}$ 中任取两个元素 γ_1,γ_2,则存在 $\alpha_1,\alpha_2\in V$,使得 $\gamma_1=\mathcal{A}\alpha_1$, $\gamma_2=\mathcal{A}\alpha_2$,从而

$$\gamma_1+\gamma_2=\mathcal{A}\alpha_1+\mathcal{A}\alpha_2=\mathcal{A}(\alpha_1+\alpha_2)\in\mathrm{Im}\,\mathcal{A},$$
$$k\gamma_1=k\mathcal{A}\alpha_1=\mathcal{A}(k\alpha_1)\in\mathrm{Im}\,\mathcal{A},\quad \forall k\in F,$$

因此 $\mathrm{Im}\,\mathcal{A}$ 是 V'的一个子空间. ∎

命题 2 设 $\mathcal{A}$ 是域 F 上线性空间 V 到 V'的一个线性映射,则

(i) $\mathcal{A}$ 是单射当且仅当 $\mathrm{Ker}\,\mathcal{A}=0$;

(ii) $\mathcal{A}$ 是满射当且仅当 $\mathrm{Im}\,\mathcal{A}=V'$.

证明 (i) 必要性.设 $\mathcal{A}$ 是单射.任取 $\alpha\in \mathrm{Ker}\,\mathcal{A}$,则

$$\mathcal{A}(\alpha)=0=\mathcal{A}(0).$$

由此推出 $\alpha=0$,因此 $\mathrm{Ker}\,\mathcal{A}=0$.

充分性.设 $\mathrm{Ker}\,\mathcal{A}=0$.设 $\alpha_1,\alpha_2\in V$,且 $\mathcal{A}\alpha_1=\mathcal{A}\alpha_2$.则

$$0=\mathcal{A}\alpha_2-\mathcal{A}\alpha_1=\mathcal{A}(\alpha_2-\alpha_1),$$

从而 $\alpha_2-\alpha_1\in \mathrm{Ker}\,\mathcal{A}$.由于 $\mathrm{Ker}\,\mathcal{A}=0$,因此 $\alpha_2-\alpha_1=0$.即 $\alpha_1=\alpha_2$.这表明 $\mathcal{A}$ 是单射.

(ii) 由满射的定义立即得到. ∎

几何空间 V 在 xOy 平面上的正投影 $\mathcal{P}_W$ 的核是 z 轴, $\mathcal{P}_W$ 的象是 xOy 平面.显然有

$$\dim\,\mathrm{Ker}\,\mathcal{P}_W+\dim\,\mathrm{Im}\,\mathcal{P}_W=1+2=3=\dim V.$$

一般地,有下述结论:

定理 3 设 V 和 V'都是域 F 上的线性空间,且 V 是有限维的,设 $\mathcal{A}$ 是 V 到 V'的一个线性映射,则 $\mathrm{Ker}\,\mathcal{A}$,$\mathrm{Im}\,\mathcal{A}$ 都是有限维的,且

$$\dim\,\mathrm{Ker}\,\mathcal{A}+\dim\,\mathrm{Im}\,\mathcal{A}=\dim V. \tag{2}$$

证明 因为 V 是有限维的,所以它的子空间 $\mathrm{Ker}\,\mathcal{A}$ 是有限维的.取 $\mathrm{Ker}\,\mathcal{A}$ 的一个基 $\alpha_1,\cdots,\alpha_m$,把它扩充成 V 的一个基

$$\alpha_1,\cdots,\alpha_m,\alpha_{m+1},\cdots,\alpha_n.$$

在 $\mathrm{Im}\,\mathcal{A}$ 中任取一个向量 $\mathcal{A}\alpha$,其中 $\alpha\in V$.设 $\alpha=\sum\limits_{i=1}^{n}a_i\alpha_i$,则

$$\mathcal{A}\alpha=\sum_{i=1}^{n}a_i\mathcal{A}\alpha_i=a_{m+1}\mathcal{A}\alpha_{m+1}+\cdots+a_n\mathcal{A}\alpha_n, \tag{3}$$

因此
$$\mathrm{Im}\,\mathcal{A}=\langle\mathcal{A}\alpha_{m+1},\cdots,\mathcal{A}\alpha_n\rangle, \tag{4}$$

从而 Im A 是有限维的.我们来证 $A\alpha_{m+1},\cdots,A\alpha_n$ 线性无关:

假设
$$k_{m+1}A\alpha_{m+1}+\cdots+k_nA\alpha_n=0,$$
则
$$A(k_{m+1}\alpha_{m+1}+\cdots+k_n\alpha_n)=0,$$
于是
$$k_{m+1}\alpha_{m+1}+\cdots+k_n\alpha_n\in \operatorname{Ker}A,$$
所以
$$k_{m+1}\alpha_{m+1}+\cdots+k_n\alpha_n=l_1\alpha_1+\cdots+l_m\alpha_m,$$
即
$$-l_1\alpha_1-\cdots-l_m\alpha_m+k_{m+1}\alpha_{m+1}+\cdots+k_n\alpha_n=0,$$
由此即得
$$l_1=\cdots=l_m=k_{m+1}=\cdots=k_n=0.$$
所以 $A\alpha_{m+1},\cdots,A\alpha_n$ 线性无关,从而它是 Im A 的一个基.因此,
$$\dim \operatorname{Im}A=n-m=\dim V-\dim \operatorname{Ker}A.$$
x

从(3)式看到,当 V 是有限维时,若 $\alpha_1,\cdots,\alpha_n$ 是 V 的一个基,

则
$$\operatorname{Im}A=\langle A\alpha_1,\cdots,A\alpha_n\rangle. \tag{5}$$

当 V 是有限维时,V 到 V'的线性映射 A 的核的维数也称为 A 的零度,A 的象 Im A 的维数称为 A 的秩.

推论 4　如果 V 和 V'都是域 F 上 n 维线性空间,且 A 是 V 到 V'的一个线性映射,则 A 是单射当且仅当 A 是满射.

证明
$$A \text{ 是单射}\quad \operatorname{Ker}A=0$$
$$\dim\operatorname{Im}A=\dim V=\dim V'$$
$$\operatorname{Im}A=V'\quad A\text{ 是满射}.$$
x

推论 5　对于有限维线性空间 V 上的线性变换 A,它是单射当且仅当它是满射.

x

注意,对于有限维线性空间 V 上的线性变换 A,虽然子空间 Ker A 与 Im A 的维数之和等于 dim V,但是 Ker A + Im A 并不一定是整个空间 V.例如,在线性空间 $K[x]_n$ 中,求导数 D 的象 $\operatorname{Im}D=K[x]_{n-1}$,$D$ 的核 $\operatorname{Ker}D=K$.显然
$$K+K[x]_{n-1}\neq K[x]_n.$$

设 A 是域 F 上一个 $s\times n$ 矩阵,令
$$A:F^n\quad F^s$$
$$\alpha\longmapsto A\alpha,$$
则 A 是 F^n 到 F^s 的一个线性映射(据§1 的例 4).在本套教材上册第 4 章§7 中,我们已证(对于 F 是数域的情形):Ker A 等于齐次线性方程组 $AX=0$ 的解空间;Im A 等于矩阵 A 的列空间.这个结论对于任意域 F 也成立.

习　题　9.2

1. 设 V 是域 F 上的一个线性空间,U,W 是 V 的两个子空间,且 $V=U\oplus W$,用 P_U 表

示平行于 W 在 U 上的投影.证明:$\operatorname{Ker}P_U = W$, $\operatorname{Im}P_U = U$.

2. 设 V 是域 F 上的一个线性空间, A 是 V 上的一个线性变换.证明:如果 A 是 V 上的幂等变换,则

$$V = \operatorname{Ker}A \oplus \operatorname{Im}A,$$

并且 A 是平行于 $\operatorname{Ker}A$ 在 $\operatorname{Im}A$ 上的投影.

3. 设 V, U, W 都是域 F 上的线性空间,并且 V 是有限维的.设 A 是 V 到 U 的一个线性映射, B 是 U 到 W 的一个线性映射.证明:

$$\dim \operatorname{Ker}(BA) \leq \dim \operatorname{Ker}A + \dim \operatorname{Ker}B.$$

4. 设 V, U, W 都是域 F 上的线性空间,并且 $\dim V = n$, $\dim U = m$,设 A 是 V 到 U 的一个线性映射, B 是 U 到 W 的一个线性映射.证明:

$$\operatorname{rank}(BA) \geq \operatorname{rank}A + \operatorname{rank}B - m.$$

5. 设 A, B 都是域 F 上线性空间 V 上的幂等变换,证明:

(1) A 与 B 有相同的象当且仅当 $AB = B$, $BA = A$;

(2) A 与 B 有相同的核当且仅当 $AB = A$, $BA = B$.

6. 设 V 和 V' 都是域 F 上的有限维线性空间, A 是 V 到 V' 的一个线性映射.证明:存在直和分解 $V = U \oplus W$, $V' = M \oplus N$,使得 $\operatorname{Ker}A = U$,且 $W \cong M$.

*7. 设 V 是域 F 上的一个线性空间, $\operatorname{char}F = 0$.证明:如果 $A_1, A_2, \cdots, A_s$ 是 V 上 s 个两两不同的线性变换,那么 V 中至少有一个向量 α,使得 $A_1\alpha, A_2\alpha, \cdots, A_s\alpha$ 两两不同.

§3 线性映射的矩阵表示

几何空间 V 在 xOy 平面上的正投影 P_W 是 V 上的一个线性变换.设 x 轴、y 轴、z 轴上的单位向量分别为 e_1, e_2, e_3.则

$$P_W(e_1) = e_1,\ P_W(e_2) = e_2,\ P_W(e_3) = 0.$$

由此得出

$$(P_W(e_1), P_W(e_2), P_W(e_3)) = (e_1, e_2, e_3)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

上式右端的 3 级矩阵刻画了正投影 P_W.由此受到启发,我们可以利用矩阵来研究线性变换.

设 V 是域 F 上的 n 维线性空间, A 是 V 上的一个线性变换.我们知道, A 被它在 V 的一个基上的作用决定.于是取 V 的一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$.由于 $A\alpha_i \in V$,因此 $A\alpha_i$ 可以被 V 的这个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 唯一地线性表出:

$$\begin{aligned}
\mathcal{A}\alpha_1 &= a_{11}\alpha_1 + a_{21}\alpha_2 + \cdots + a_{n1}\alpha_n,\\
\mathcal{A}\alpha_2 &= a_{12}\alpha_1 + a_{22}\alpha_2 + \cdots + a_{n2}\alpha_n,\\
&\cdots\cdots\cdots\cdots\\
\mathcal{A}\alpha_n &= a_{1n}\alpha_1 + a_{2n}\alpha_2 + \cdots + a_{nn}\alpha_n.
\end{aligned} \tag{1}$$

按照形式写法,(1)式可以写成

$$(\mathcal{A}\alpha_1, \mathcal{A}\alpha_2, \cdots, \mathcal{A}\alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & & \cdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}. \tag{2}$$

我们把(2)式右端的 n 级矩阵(a_{ij})记作 A,把 A 称为线性变换 $\mathcal{A}$ 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵. A 的第 j 列是 $\mathcal{A}\alpha_j$ 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标, $j = 1, 2, \cdots, n$.因此 A 被线性变换 $\mathcal{A}$ 唯一决定.

我们把$(\mathcal{A}\alpha_1, \mathcal{A}\alpha_2, \cdots, \mathcal{A}\alpha_n)$简记成 $\mathcal{A}(\alpha_1, \alpha_2, \cdots, \alpha_n)$.于是从(2)式得, n 级矩阵 A 是 V 上线性变换 $\mathcal{A}$ 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵当且仅当下式成立:

$$\mathcal{A}(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n) A. \tag{3}$$

例1　在 $\mathbf{R}^{\mathbf{R}}$ 中,设

$$V = \langle 1, \sin x, \cos x\rangle.$$

求导数 $\mathcal{D}$ 是 V 上的线性变换(因为 $\mathcal{D}(k_1\cdot 1 + k_2\sin x + k_3\cos x) = k_2\cos x - k_3\sin x \in V$);$1, \sin x, \cos x$ 是 V 的一个基(因为它们线性无关).写出 $\mathcal{D}$ 在基 $1, \sin x, \sin x$ 下的矩阵.

解　因为

$$\begin{aligned}
\mathcal{D}(1) &= 0 = 0\cdot 1 + 0\cdot\sin x + 0\cdot\cos x,\\
\mathcal{D}(\sin x) &= \cos x = 0\cdot 1 + 0\cdot\sin x + 1\cdot\cos x,\\
\mathcal{D}(\cos x) &= -\sin x = 0\cdot 1 + (-1)\sin x + 0\cdot\cos x,
\end{aligned}$$

所在 $\mathcal{D}$ 在基 $1, \sin x, \cos x$ 下的矩阵是

$$D = \begin{pmatrix} 0 & 0 & 0\\ 0 & 0 & -1\\ 0 & 1 & 0\end{pmatrix}.$$

上述说明, n 维线性空间 V 上的线性变换可以用矩阵来表示.下面我们来讨论两个有限维线性空间之间的线性映射能不能用矩阵来表示.

设 V 和 V'分别是域 F 上 n 维、s 维线性空间, $\mathcal{A}$ 是 V 到 V'的一个线性映射.在 V 中取一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$;在 V'中取一个基 $\eta_1, \eta_2, \cdots, \eta_s$.由于 $\mathcal{A}\alpha_i \in V'$,因此 $\mathcal{A}\alpha_i$ 可以由 V'的基 $\eta_1, \eta_2, \cdots, \eta_s$ 唯一地线性表出:

$$\begin{aligned}\mathcal{A}\alpha_1 &= a_{11}\eta_1 + a_{21}\eta_2 + \cdots + a_{s1}\eta_s,\\ \mathcal{A}\alpha_2 &= a_{12}\eta_1 + a_{22}\eta_2 + \cdots + a_{s2}\eta_s,\\ &\cdots\cdots\cdots\cdots\\ \mathcal{A}\alpha_n &= a_{1n}\eta_1 + a_{2n}\eta_2 + \cdots + a_{sn}\eta_s.\end{aligned} \tag{4}$$

按照形式写法,(4)式可以写成

$$(\mathcal{A}\alpha_1, \mathcal{A}\alpha_2, \cdots, \mathcal{A}\alpha_n) = (\eta_1, \eta_2, \cdots, \eta_s)\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & & \cdots\\ a_{s1} & a_{s2} & \cdots & a_{sn}\end{pmatrix}. \tag{5}$$

我们把(5)式右端的 $s\times n$ 矩阵(a_{ij})记作 A,称 A 是线性映射 $\mathcal{A}$ 在 V 的基 α_1, $\alpha_2,\cdots,\alpha_n$ 和 V'的基 $\eta_1,\eta_2,\cdots,\eta_s$ 下的矩阵. A 的第 j 列是 $\mathcal{A}\alpha_j$ 在基 $\eta_1,\eta_2,\cdots,$ η_s 下的坐标,$j=1,2,\cdots,n$,因此 A 被线性映射 $\mathcal{A}$ 唯一决定.从(5)式得 $s\times n$ 矩阵 A 是 V 到 V'的线性映射 $\mathcal{A}$ 在 V 的基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 和 V'的基 $\eta_1,\eta_2,\cdots,\eta_s$ 下的矩阵当且仅当下式成立:

$$\mathcal{A}(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\eta_1,\eta_2,\cdots,\eta_s)A. \tag{6}$$

从上面看到,域 F 上 n 维线性空间 V 到 s 维线性空间 V'的每一个线性映射 $\mathcal{A}$ 可以用一个 $s\times n$ 矩阵 A 来表示.我们在本章第 1 节已经知道,V 到 V'的所有线性映射组成的集合 $\mathrm{Hom}(V,V')$是域 F 上一个线性空间.我们又知道,F 上所有 $s\times n$ 矩阵组成的集合 $M_{s\times n}(F)$也是域 F 上一个线性空间.试问:这两个线性空间之间有什么关系?

在 V 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$;在 V'中取一个基 $\eta_1,\eta_2,\cdots,\eta_s$,令

$$\begin{aligned}\sigma:\ \mathrm{Hom}(V,V') &\to M_{s\times n}(F)\\ \mathcal{A} &\longmapsto A,\end{aligned}$$

其中 A 是线性映射 $\mathcal{A}$ 在 V 的基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 和 V'的基 $\eta_1,\eta_2,\cdots,\eta_s$ 下的矩阵.

显然,σ 是 $\mathrm{Hom}(V,V')$到 $M_{s\times n}(F)$的一个映射.先看 σ 是否单射?设 V 到 V'的线性映射 $\mathcal{B}$ 在 V 的基 $\alpha_1,\cdots,\alpha_n$ 和 V'的基 $\eta_1,\cdots,\eta_s$ 下的矩阵为 B,则 $\sigma(\mathcal{B})=B$.又 $\sigma(\mathcal{A})=A$.如果 $A=B$,由于 A 的第 j 列是 $\mathcal{A}\alpha_j$ 在 V'的基 η_1, $\eta_2,\cdots,\eta_s$ 下的坐标,B 的第 j 列是 $\mathcal{B}\alpha_j$ 在这个基下的坐标,因此 $\mathcal{A}\alpha_j=\mathcal{B}\alpha_j$,$j=1,2,\cdots,n$.从而对一切 $\alpha\in V$,有 $\mathcal{A}\alpha=\mathcal{B}\alpha$.因此 $\mathcal{A}=\mathcal{B}$,这表明 σ 是单射.

再看 σ 是否满射.任给 $C\in M_{s\times n}(F)$,令

$$(\gamma_1,\gamma_2,\cdots,\gamma_n)=(\eta_1,\eta_2,\cdots,\eta_s)C, \tag{7}$$

显然 $\gamma_1,\gamma_2,\cdots,\gamma_n\in V'$.据本章§1 的定理 1,存在 V 到 V'的一个线性映射 $\mathcal{C}$,

使得 $\mathcal{C}\alpha_j = \gamma_j$, $j = 1,2,\cdots,n$,于是

$$\mathcal{C}(\alpha_1,\alpha_2,\cdots,\alpha_n) = (\gamma_1,\gamma_2,\cdots,\gamma_n) = (\eta_1,\eta_2,\cdots,\eta_s)C. \tag{8}$$

(8)式表明, C 是线性映射 $\mathcal{C}$ 的矩阵,因此 $\sigma(\mathcal{C}) = C$.这表明 σ 是满射.

上述表明 σ 是 $\mathrm{Hom}(V,V')$到 $M_{s\times n}(F)$的一个双射.现在来看 σ 是否保持加法与数量乘法运算.

设 $\mathcal{A},\mathcal{B}\in \mathrm{Hom}(V,V')$, $k\in F$.设 $\sigma(\mathcal{A}) = A$, $\sigma(\mathcal{B}) = B$.由于

$$\begin{aligned}
&(\mathcal{A}+\mathcal{B})(\alpha_1,\alpha_2,\cdots,\alpha_n)\\
&= (\mathcal{A}\alpha_1+\mathcal{B}\alpha_1,\ \mathcal{A}\alpha_2+\mathcal{B}\alpha_2,\cdots,\ \mathcal{A}\alpha_n+\mathcal{B}\alpha_n)\\
&= (\mathcal{A}\alpha_1,\mathcal{A}\alpha_2,\cdots,\mathcal{A}\alpha_n)+(\mathcal{B}\alpha_1,\mathcal{B}\alpha_2,\cdots,\mathcal{B}\alpha_n)\\
&= (\eta_1,\eta_2,\cdots,\eta_s)A+(\eta_1,\eta_2,\cdots,\eta_s)B\\
&= (\eta_1,\eta_2,\cdots,\eta_s)(A+B),
\end{aligned}$$

因此线性映射 $\mathcal{A}+\mathcal{B}$ 的矩阵是 $A+B$.从而

$$\sigma(\mathcal{A}+\mathcal{B}) = A+B = \sigma(\mathcal{A})+\sigma(\mathcal{B}),$$

这表明 σ 保持加法运算.我们又有

$$\begin{aligned}
&(k\mathcal{A})(\alpha_1,\alpha_2,\cdots,\alpha_n)\\
&= (k\mathcal{A}\alpha_1,k\mathcal{A}\alpha_2,\cdots,k\mathcal{A}\alpha_n)\\
&= k(\mathcal{A}\alpha_1,\mathcal{A}\alpha_2,\cdots,\mathcal{A}\alpha_n) = k(\eta_1,\eta_2,\cdots,\eta_s)A\\
&= (\eta_1,\eta_2,\cdots,\eta_s)(kA),
\end{aligned}$$

因此线性映射 $k\mathcal{A}$ 的矩阵是 kA.从而

$$\sigma(k\mathcal{A}) = kA = k\sigma(\mathcal{A}), \tag{9}$$

这表明 σ 保持数量乘法运算.

综上述, σ 是线性空间 $\mathrm{Hom}(V,V')$到 $M_{s\times n}(F)$的一个同构映射,因此我们得到下述结论:

定理 1　设 V 和 V'分别是域 F 上 n 维、s 维线性空间,则

$$\mathrm{Hom}(V,V')\cong M_{s\times n}(F) \tag{10}$$

从而

$$\dim \mathrm{Hom}(V,V') = \dim M_{s\times n}(F) = sn = (\dim V)(\dim V'). \tag{11}$$

推论 2　设 V 是域 F 上的 n 维线性空间,则

$$\mathrm{Hom}(V,V)\cong M_n(F), \tag{12}$$

且

$$\dim \mathrm{Hom}(V,V) = (\dim V)^2. \tag{13}$$

在 $\mathrm{Hom}(V,V)$与 $M_n(F)$中,都有乘法运算.我们可以进一步证明:把线性变换 $\mathcal{A}$ 对应到它在 V 的基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵 A 的映射 σ 还保持乘法运算:设线性变换 $\mathcal{B}$ 在 V 的基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵是 $B=(b_{ij})$.由于

$$
\begin{aligned}
&(AB)(\alpha_1,\alpha_2,\cdots,\alpha_n)\\
&= A(B\alpha_1, B\alpha_2,\cdots, B\alpha_n)\\
&= A\,(\alpha_1,\alpha_2,\cdots,\alpha_n)B\\
&= A\,(b_{11}\alpha_1 + b_{21}\alpha_2 + \cdots + b_{n1}\alpha_n,\cdots, b_{1n}\alpha_1 + b_{2n}\alpha_2 + \cdots + b_{nn}\alpha_n)\\
&= (b_{11}A\alpha_1 + b_{21}A\alpha_2 + \cdots + b_{n1}A\alpha_n,\cdots, b_{1n}A\alpha_1 + b_{2n}A\alpha_2 + \cdots + b_{nn}A\alpha_n)\\
&= (A\alpha_1, A\alpha_2,\cdots, A\alpha_n)\begin{pmatrix} b_{11} & \cdots & b_{1n}\\ \cdots & & \cdots\\ b_{n1} & \cdots & b_{nn}\end{pmatrix}\\
&= A(\alpha_1,\alpha_2,\cdots,\alpha_n)\,B\\
&= (\alpha_1,\alpha_2,\cdots,\alpha_n)A\,B = (\alpha_1,\alpha_2,\cdots,\alpha_n)(AB),
\end{aligned}
\tag{14}
$$

因此 AB 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵是 AB.从而

$$\sigma(AB) = AB = \sigma(A)\sigma(B), \tag{15}$$

这表明 σ 保持乘法运算.

从(14)式的推导过程中还可看到,

$$A\,(\alpha_1,\alpha_2,\cdots,\alpha_n)B = A(\alpha_1,\alpha_2,\cdots,\alpha_n)\,B. \tag{16}$$

显然, V 上的恒等变换 $\mathbf{I}$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵是单位矩阵 $\mathbf{I}$, 因此 $\sigma(\mathbf{I}) = \mathbf{I}$.

设 V 上线性变换 A 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵是 A, 由于

线性变换 A 可逆

存在 V 上的线性变换 B 使得 $AB = BA = \mathbf{I}$

存在 V 上的线性变换 B 使得 $\sigma(A)\sigma(B) = \sigma(B)\sigma(A) = \sigma(\mathbf{I})$

存在域 F 上 n 级矩阵 B 使得 $AB = BA = \mathbf{I}$

矩阵 A 可逆,

因此, V 上线性变换 A 可逆当且仅当它在 V 的一个基下的矩阵 A 可逆.从上述推导过程还可看到,设线性变换 A, B 在 V 的一个基下的矩阵分别为 A, B,则 B 是可逆线性变换 A 的逆变换当且仅当 B 是可逆矩阵 A 的逆矩阵.

设 A 是域 F 上 n 维线性空间 V 的一个线性变换,且 A 在 V 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵是 A. V 中任一向量 α 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标记作 X, 问: $A\alpha$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标是什么?

由于 $\alpha = (\alpha_1,\alpha_2,\cdots,\alpha_n)X$,因此利用(16)式,得到

$$
\begin{aligned}
A\alpha &= A\,(\alpha_1,\alpha_2,\cdots,\alpha_n)X = A(\alpha_1,\alpha_2,\cdots,\alpha_n)\,X\\
&= (\alpha_1,\alpha_2,\cdots,\alpha_n)A\,X = (\alpha_1,\alpha_2,\cdots,\alpha_n)(AX).
\end{aligned}
$$

这表明 $A\alpha$ 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标是 AX.

由于 V 中两个向量相等当且仅当它们在 V 的一个基下的坐标相等,因此,如果向量 γ 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标为 Y,则

$$A\alpha = \gamma \qquad AX = Y. \tag{17}$$

域 F 上 n 维线性空间 V 上的一个线性变换 A 在 V 的不同基下的矩阵有什么关系?

定理 3　设 V 是域 F 上 n 维线性空间, V 上的一个线性变换 A 在 V 的两个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 与 $\eta_1, \eta_2, \cdots, \eta_n$ 下的矩阵分别为 A, B. 从基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 到基 $\eta_1, \eta_2, \cdots, \eta_n$ 的过渡矩阵是 S,则

$$B = S^{-1}AS. \tag{18}$$

证明　由已知条件,我们有

$$A(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)A,$$
$$A(\eta_1, \eta_2, \cdots, \eta_n) = (\eta_1, \eta_2, \cdots, \eta_n)B,$$
$$(\eta_1, \eta_2, \cdots, \eta_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)S,$$

于是

$$\begin{aligned}(\eta_1, \eta_2, \cdots, \eta_n)S^{-1} &= (\alpha_1, \alpha_2, \cdots, \alpha_n)S\ S^{-1}\\ &= (\alpha_1, \alpha_2, \cdots, \alpha_n)(SS^{-1}) = (\alpha_1, \alpha_2, \cdots, \alpha_n),\end{aligned}$$

从而

$$\begin{aligned}&A(\eta_1, \eta_2, \cdots, \eta_n)\\ &= A\ (\alpha_1, \alpha_2, \cdots, \alpha_n)S\\ &= \ A(\alpha_1, \alpha_2, \cdots, \alpha_n)\ S = (\alpha_1, \alpha_2, \cdots, \alpha_n)A\ S\\ &= (\alpha_1, \alpha_2, \cdots, \alpha_n)(AS) = (\eta_1, \eta_2, \cdots, \eta_n)S^{-1}\ (AS)\\ &= (\eta_1, \eta_2, \cdots, \eta_n)\ S^{-1}AS\ .\end{aligned} \tag{19}$$

(19)式表明, A 在基 $\eta_1, \eta_2, \cdots, \eta_n$ 下的矩阵是 $S^{-1}AS$. 由于 A 在基 $\eta_1, \eta_2, \cdots, \eta_n$ 下的矩阵是唯一的,因此 $B = S^{-1}AS$. x

定理 3 表明,同一个线性变换 A 在 V 的不同基下的矩阵是相似的.

习　题　9.3

1. 设 A 是 K^3 上的一个线性变换:

$$A\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}x_1 + 2x_2\\x_3 - x_2\\x_2 - x_3\end{pmatrix},$$

求 A 在标准基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 下的矩阵.

2. 在 R^R 中,由下述两个函数

$$f_1 = e^{ax}\cos bx,\ f_2 = e^{ax}\sin bx$$

生成的 2 维子空间记作 V,说明求导数 D 是 V 上的一个线性变换,并且求 D 在 V 的一个基 f_1, f_2 下的矩阵.

3. $\mathcal{A}$ 是 $M_2(K)$上的一个线性变换:

$$\mathcal{A}(X) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} X, \quad \forall X \in M_2(K),$$

求 $\mathcal{A}$ 在基 $E_{11}, E_{12}, E_{21}, E_{22}$ 下的矩阵.

4. $\mathcal{B}$ 是 $M_2(K)$上的一个线性变换:

$$\mathcal{B}(X) = X \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \forall X \in M_2(K),$$

求 $\mathcal{B}$ 在基 $E_{11}, E_{12}, E_{21}, E_{22}$ 下的矩阵.

5. 设 V 是域 F 上的 n 维线性空间,设有 V 上的线性变换 $\mathcal{A}$ 与 V 中的向量 α,使得 $\mathcal{A}^{n-1}\alpha \neq 0$,且 $\mathcal{A}^n\alpha = 0$.证明: V 中存在一个基,使得 $\mathcal{A}$ 在这个基下的矩阵是

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

6. 设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的一个线性变换, $\mathcal{A}$ 在 V 的一个基下的矩阵是 A.证明: $\mathcal{A}$ 是幂等变换当且仅当 A 是幂等矩阵.

7. 设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的一个线性变换.证明:

(1) 在 $F[x]$中存在一个次数 $\leq n^2$ 的非零多项式 $f(x)$使得 $f(\mathcal{A}) = 0$;

(2) 如果 $f(\mathcal{A}) = 0, g(\mathcal{A}) = 0$,则 $d(\mathcal{A}) = 0$,这里 $d(x)$是 $f(x)$与 $g(x)$的一个最大公因式;

(3) $\mathcal{A}$ 可逆的充分必要条件是:有一个常数项不为零的多项式 $f(x)$使得 $f(\mathcal{A}) = 0$.

8. 设 V 是域 F 上的 n 维线性空间, $\alpha_1, \cdots, \alpha_n$ 是 V 的一个基.设 V 上的一个线性变换 $\mathcal{A}$ 在基 $\alpha_1, \cdots, \alpha_n$ 下的矩阵是 A,证明:

$$\text{rank}\ \mathcal{A} = \text{rank}\ A$$

*9. 设 V 和 V' 都是域 F 上有限维线性空间, $\mathcal{A}$ 是 V 到 V'的一个线性映射.证明:存在 V 的一个基和 V' 的一个基,使得 $\mathcal{A}$ 在这一对基下的矩阵 A 形如

$$\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

10. 已知 K^3 上的线性变换 $\mathcal{A}$ 在标准基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 下的矩阵是

$$A = \begin{pmatrix} 15 & -11 & 5 \\ 20 & -15 & 8 \\ 8 & -7 & 6 \end{pmatrix},$$

设 $\eta_1 = (2,3,1)', \eta_2 = (3,4,1)', \eta_3 = (1,2,2)'$,求 $\mathcal{A}$ 在基 η_1, η_2, η_3 下的矩阵 B.

11. 已知 K^3 上线性变换 $\mathcal{A}$ 在基 $\alpha_1 = (8,-6,7)', \alpha_2 = (-16,7,-13)', \alpha_3 = (9,-3,7)'$

下的矩阵为

$$A=\begin{pmatrix}1&-18&15\\-1&-22&20\\1&-25&22\end{pmatrix},$$

求 $\mathcal{A}$ 在基 $\eta_1=(1,-2,1)',\eta_2=(3,-1,2)',\eta_3=(2,1,2)'$下的矩阵 B.

12. 设域 F 上 3 维线性空间 V 上的线性变换 $\mathcal{A}$ 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的矩阵为

$$A=\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix},$$

(1) 求 $\mathcal{A}$ 在基 $\alpha_2,\alpha_3,\alpha_1$ 下的矩阵;

(2) 求 $\mathcal{A}$ 在基 $k\alpha_1,\alpha_2,\alpha_3$ 下的矩阵,其中 $k\in F$ 且 $k\neq 0$;

(3) 求 $\mathcal{A}$ 在基 $\alpha_1,\alpha_1+\alpha_2,\alpha_3$ 下的矩阵.

*13. 设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的一个线性变换.证明:如果 $\mathcal{A}$ 在 V 的所有各个基下的矩阵都相同,则 $\mathcal{A}$ 是数乘变换.

14. 设 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 是数域 K 上 4 维线性空间 V 的一个基,已知线性变换 $\mathcal{A}$ 在这个基下的矩阵为

$$A=\begin{pmatrix}1&0&2&1\\-1&2&1&3\\1&2&5&5\\2&-2&1&-2\end{pmatrix},$$

(1) 求 $\mathcal{A}$ 在基 $\eta_1=\alpha_1-2\alpha_2+\alpha_4,\eta_2=3\alpha_2-\alpha_3-\alpha_4,\eta_3=\alpha_3+\alpha_4,\eta_4=2\alpha_4$ 下的矩阵;

(2) 求 $\mathcal{A}$ 的核与值域;

(3) 在 $\operatorname{Ker}\mathcal{A}$ 中选一个基,把它扩充成 V 的一个基,并求 $\mathcal{A}$ 在这个基下的矩阵;

(4) 在 $\operatorname{Im}\mathcal{A}$ 中选一个基,把它扩充成 V 的一个基,并且求 $\mathcal{A}$ 在这个基下的矩阵.

§4 线性变换的特征值与特征向量

上一节的定理 3 表明,域 F 上 n 维线性空间 V 上的线性变换 $\mathcal{A}$ 在 V 的不同基下的矩阵是相似的.由于相似的矩阵有相同的行列式、秩、迹、特征多项式、特征值,因此我们可以把线性变换 $\mathcal{A}$ 在 V 的一个基下的矩阵 A 的特征值、秩、迹、特征多项式、特征值分别叫做线性变换 $\mathcal{A}$ 的行列式、秩、迹、特征多项式、特征值(在§2 中,线性变换 $\mathcal{A}$ 的秩定义成$\operatorname{Im}\mathcal{A}$ 的维数,现在又把 $\mathcal{A}$ 在 V 的一个基下的矩阵 A 的秩叫做 $\mathcal{A}$ 的秩,这是因为在习题 9.3 的第 8 题已证明 $\dim\operatorname{Im}\mathcal{A}=\operatorname{rank}A$).

为了更好地理解线性变换的特征值的几何意义,以及对无限维线性空间上的线性变换也考虑它的特征值,我们给出如下的定义:

定义 1　设 A 是域 F 上线性空间 V 上的一个线性变换，如果 V 中存在一个非零向量 ξ，使得

$$A\xi = \lambda_0 \xi, \quad \lambda_0 \in F \tag{1}$$

则称 λ_0 是 A 的一个特征值，称 ξ 是 A 的属于特征值 λ_0 的一个特征向量.

从定义 1 看出，线性变换 A 的特征向量 ξ 有这样的"几何意义"：A 对 ξ 的作用是把 ξ"拉伸(或压缩)"λ_0 倍，这个倍数 λ_0 就是 A 的一个特征值.

现在设 V 是域 F 上 n 维线性空间，V 中取定一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. V 上的一个线性变换 A 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵是 A，向量 ξ 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标是 X，$\lambda_0 \in F$. 从§3 的(17)式得出

$$A\xi = \lambda_0 \xi \qquad AX = \lambda_0 X. \tag{2}$$

由此得出

$$\lambda_0 \text{是 A 的一个特征值} \qquad \lambda_0 \text{是 A 的一个特征值}; \tag{3}$$

ξ 是 A 的属于特征值 λ_0 的一个特征向量

ξ 的坐标 X 是 A 的属于特征值 λ_0 的一个特征向量.　(4)

(3)式说明，对于有限维线性空间，用线性变换的矩阵的特征值定义成线性变换的特征值，与定义 1 是一致的.

(3)式和(4)式给出了求有限维线性空间上的线性变换 A 的全部特征值和特征向量的方法：只要去求 A 在 V 的一个基下的矩阵 A 的全部特征值和特征向量. 但是要注意：矩阵 A 的特征向量 X 是线性变换 A 的特征向量 ξ 的坐标.

例 1　设 V 是数域 K 上 3 维线性空间，A 是 V 上一个线性变换，A 在 V 的一个基 $\alpha_1, \alpha_2, \alpha_3$ 下的矩阵是

$$A = \begin{matrix} 2 & -2 & 2 \\ -2 & -1 & 4 \\ 2 & 4 & -1 \end{matrix},$$

求 A 的全部特征值和特征向量.

解　A 的特征多项式为

$$|\lambda I - A| = \begin{vmatrix} \lambda - 2 & 2 & -2 \\ 2 & \lambda + 1 & -4 \\ -2 & -4 & \lambda + 1 \end{vmatrix} = (\lambda - 3)^2 (\lambda + 6),$$

于是 A 的全部特征值是 3(二重)，-6.

对于特征值 3，解齐次线性方程组 $(3I - A)X = 0$，得到一个基础解系：

$$\begin{matrix} -2 \\ 1 \\ 0 \end{matrix}, \quad \begin{matrix} 2 \\ 0 \\ 1 \end{matrix}.$$

令
$$\xi_1 = -2\alpha_1 + \alpha_2,\ \xi_2 = 2\alpha_1 + \alpha_3,$$
则 $\mathcal{A}$ 的属于特征值 3 的全部特征向量是
$$\{k_1\xi_1 + k_2\xi_2 \mid k_1, k_2 \in K, \text{且 } k_1, k_2 \text{ 不全为 } 0\}.$$

对于特征值 -6, 求出 $(-6I - A)X = 0$ 的一个基础解系:
$$\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}.$$
令
$$\xi_3 = \alpha_1 + 2\alpha_2 - 2\alpha_3,$$
则 $\mathcal{A}$ 的属于特征值 -6 的全部特征向量是
$$\{k\xi_3 \mid k \in K, \text{且 } k \neq 0\}.$$

设 $\mathcal{A}$ 是域 F 上线性空间 V 上的一个线性变换, λ_0 是 $\mathcal{A}$ 的一个特征值. 令
$$V_{\lambda_0} \overset{\text{def}}{=} \{\alpha \mid \mathcal{A}\alpha = \lambda_0\alpha, \alpha \in V\}, \tag{5}$$
则易验证 V_{λ_0} 是 V 的一个子空间, 称 V_{λ_0} 是 $\mathcal{A}$ 的属于特征值 λ_0 的特征子空间. V_{λ_0} 中的全部非零向量就是 $\mathcal{A}$ 的属于 λ_0 的全部特征值.

设 V 是域 F 上 n 维线性空间, V 上线性变换 $\mathcal{A}$ 在 V 的一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵为 A, λ_0 是 $\mathcal{A}$ 的一个特征值, 则从(5)式和(2)式得出, 线性变换 $\mathcal{A}$ 的属于 λ_0 的特征子空间 V_{λ_0} 的所有向量的坐标组成的集合是矩阵 A 的属于 λ_0 的特征子空间, 后者就是齐次线性方程组 $(\lambda_0 I - A)X = 0$ 的解空间, 因此
$$\dim V_{\lambda_0} = n - \text{rank}(\lambda_0 I - A). \tag{6}$$

无论是理论上, 还是实际应用上, 都希望对于线性变换 $\mathcal{A}$, 在 V 中能找到一个适当的基, 使得 $\mathcal{A}$ 在这个基下的矩阵具有最简单的形式. 由于 $\mathcal{A}$ 在 V 的不同基下的矩阵是相似的, 因此这个问题也就是求 $\mathcal{A}$ 在 V 的一个基下的矩阵 A 的相似标准形. 我们曾在第五章讨论过 n 级矩阵 A 的相似标准形问题, 给出了 A 的相似标准形为对角矩阵(即 A 可对角化)的充分必要条件. 但是对于不可以对角化的矩阵 A, 它的相似标准形是什么? 我们尚未进行讨论. 从现在开始, 我们将对线性变换 $\mathcal{A}$, 研究如何找 V 的一个适当的基, 使得 $\mathcal{A}$ 在这个基下的矩阵具有最简单的形式.

如果 V 中存在一个基, 使得线性变换 $\mathcal{A}$ 在这个基下的矩阵是对角矩阵, 则称 $\mathcal{A}$ 可对角化.

设线性变换 $\mathcal{A}$ 在 V 的一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵为 A, 则 $\mathcal{A}$ 可对角化当且仅当 A 可对角化. 于是从 n 级矩阵 A 可对角化的充分必要条件(见第 5 章 §6 的定理 1、定理 5, 第 8 章习题 8.2 的第 16 题), 以及 V 与 F^n 同构的性质, 可以得出线性变换 $\mathcal{A}$ 可对角化的充分必要条件如下:

定理 1 设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的一个线性变换,则

$\mathcal{A}$ 可对角化

$\mathcal{A}$ 有 n 个线性无关的特征向量

V 中存在由 $\mathcal{A}$ 的特征向量组成的一个基

$\mathcal{A}$ 的属于不同特征值的特征子空间的维数之和等于 n

$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_s}$,

其中 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是 $\mathcal{A}$ 的所有不同的特征值. x

如果 $\mathcal{A}$ 可对角化,则 $\mathcal{A}$ 有 n 个线性无关的特征向量 $\xi_1, \xi_2, \cdots, \xi_n$,其中 $\mathcal{A}\xi_i = \lambda_i \xi_i$, $i = 1, 2, \cdots, n$. 于是

$$\mathcal{A}(\xi_1, \xi_2, \cdots, \xi_n) = (\xi_1, \xi_2, \cdots, \xi_n)\begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & & \cdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}. \tag{7}$$

从(7)式看出,其右端的对角矩阵的主对角元恰好是 $\mathcal{A}$ 的全部特征值(重根按重数计算). 因此这个对角矩阵除了主对角线上元素的排列次序外,是由线性变换 $\mathcal{A}$ 唯一决定的,我们把这个对角矩阵称为线性变换 $\mathcal{A}$ 的标准形.

例 2 例 1 中的线性变换 $\mathcal{A}$ 是否可对角化? 如果 $\mathcal{A}$ 可对角化,求 $\mathcal{A}$ 的标准形.

解 在例 1 中已求出 $\mathcal{A}$ 的属于特征值 3 的两个线性无关的特征向量 ξ_1, ξ_2, $\mathcal{A}$ 的属于 -6 的一个特征向量 ξ_3. 我们从第 5 章§6 知道,矩阵 A 的属于不同特征值的特征向量是线性无关的. 利用线性空间同构的性质得出, ξ_1, ξ_2, ξ_3 也是线性无关的. 因此 $\mathcal{A}$ 可对角化.

由于 $\mathcal{A}$ 的全部特征值是 3(二重), -6,因此 $\mathcal{A}$ 的标准形是

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -6 \end{pmatrix}.$$

对于不可以对角化的线性变换 $\mathcal{A}$,能不能在 V 中找到一个适当的基,使得 $\mathcal{A}$ 在这个基下的矩阵具有最简单形式呢? 这个最简单形式当然不是对角矩阵了,那么它是什么样的矩阵呢? 这个问题在下面一节来开始讨论.

习 题 9.4

1. 设 V 是数域 K 上 3 维线性空间, V 上的一个线性变换 $\mathcal{A}$ 在 V 的一个基 $\alpha_1, \alpha_2, \alpha_3$ 下的矩阵为 A,求 $\mathcal{A}$ 的全部特征值和特征向量.

(1) $A = \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix}$;　(2) $A = \begin{pmatrix} 2 & 3 & 2 \\ 1 & 8 & 2 \\ -2 & -14 & -3 \end{pmatrix}$.

2. 第 1 题中的线性变换 $\mathscr{A}$ 是否可对角化?如果 $\mathscr{A}$ 可对角化,求 $\mathscr{A}$ 的标准形.

3. 设 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 是数域 K 上 4 维线性空间 V 的一个基,V 上的线性变换 $\mathscr{A}$ 在这个基下的矩阵为

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

(1) 求 $\mathscr{A}$ 的全部特征值与特征向量;

(2) 求 V 的一个基,使得 $\mathscr{A}$ 在这个基下的矩阵为对角矩阵,并且写出这个对角矩阵.

4. 设 V 是域 F 上任意一个线性空间(可以是无限维的),$\mathscr{A}$ 是 V 上的一个线性变换.证明:$\mathscr{A}$ 的属于不同特征值的特征向量是线性无关的.

5. 设 V 是域 F 上任意一个线性空间,λ_1, λ_2 是线性变换 $\mathscr{A}$ 的两个不同的特征值,ξ_1, ξ_2 分别是属于 λ_1, λ_2 的特征向量.证明:$\xi_1 + \xi_2$ 不是 $\mathscr{A}$ 的特征向量.

6. 设 V 是域 F 上任一线性空间,证明:如果 V 上的线性变换 $\mathscr{A}$ 以 V 中每个非零向量作为它的特征向量,则 $\mathscr{A}$ 是数乘变换.

7. 设 V 是域 F 上任一线性空间,设 $\mathscr{A}$ 是 V 上的可逆的线性变换.

(1) 证明 $\mathscr{A}$ 的特征值一定不为 0;

(2) 证明:如果 λ 是 $\mathscr{A}$ 的特征值,则 $\frac{1}{\lambda}$ 是 $\mathscr{A}^{-1}$ 的特征值.

8. 设 $\mathscr{A}$ 是域 F 上 n 维线性空间 V 上的一个线性变换,$\lambda_1, \lambda_2, \cdots, \lambda_s$ 是 $\mathscr{A}$ 的所有不同的特征值.证明:

$$\mathscr{A} \text{ 可对角化} \Leftrightarrow V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_s}.$$

*9. 设 $\mathscr{A}$ 是复数域上 n 维线性空间 V 上的一个线性变换,$\mathscr{A}$ 的所有特征值组成的 n 元数组 $(\lambda_1, \lambda_2, \cdots, \lambda_n)$ 称为 $\mathscr{A}$ 的谱.如果 $\mathscr{A}$ 的所有特征值都是 1 重的,则 $\mathscr{A}$ 的谱称为单的.证明:如果 $\mathscr{A}$ 的谱是单的,那么与 $\mathscr{A}$ 可交换的每一个线性变换 $\mathscr{B}$ 能唯一地表示成 $\mathscr{A}$ 的一个次数小于 n 的多项式.

§5　线性变换的不变子空间

上一节我们讨论了可对角化的线性变换,不可以对角化的线性变换,其结构又如何呢?解决这个问题的思路是什么?从§4 的定理 1 看到,$\mathscr{A}$ 是可对角化的线性变换当且仅当空间 V 能分解成 $\mathscr{A}$ 的特征子空间的直和.由此受到启发,研究不可以对角化的线性变换的结构,能不能从研究空间 V 分解成与 $\mathscr{A}$ 有关

的特殊类型的子空间的直和入手？这节就来讨论这个问题.注意 $\mathcal{A}$ 的特征子空间 V_{λ_i} 具有如下性质：若 $\alpha\in V_{\lambda_i}$，则 $\mathcal{A}\alpha=\lambda_i\alpha\in V_{\lambda_i}$.这启发我们引入 $\mathcal{A}$ 的不变子空间的概念.

定义 1 设 $\mathcal{A}$ 是域 F 上线性空间 V 的线性变换，W 是 V 的子空间.如果 W 中的向量在 $\mathcal{A}$ 下的象仍在 W 中，即对于任意 $\alpha\in W$，都有 $\mathcal{A}\alpha\in W$，则称 W 是 $\mathcal{A}$ 的不变子空间，简称 $\mathcal{A}$-子空间.

显然，整个空间 V 和零子空间 0，对于 V 上的每个线性变换 $\mathcal{A}$ 来说，都是 $\mathcal{A}$-子空间.称 V 和 0 是 $\mathcal{A}$ 的平凡的不变子空间.

命题 1 V 上线性变换 $\mathcal{A}$ 的核与象，$\mathcal{A}$ 的特征子空间都是 $\mathcal{A}$-子空间.

证明 任取 $\alpha\in \mathrm{Ker}\,\mathcal{A}$，因为 $\mathcal{A}\alpha=0\in\mathrm{Ker}\,\mathcal{A}$，所以 $\mathrm{Ker}\,\mathcal{A}$ 是 $\mathcal{A}$-子空间.

任取 $\alpha\in\mathrm{Im}\,\mathcal{A}$，因为 $\mathcal{A}\alpha\in\mathrm{Im}\,\mathcal{A}$，所以 $\mathrm{Im}\,\mathcal{A}$ 是 $\mathcal{A}$-子空间.

任取 $\alpha\in V_{\lambda_i}$，因为 $\mathcal{A}\alpha=\lambda_i\alpha\in V_{\lambda_i}$，所以 V_{λ_i} 是 $\mathcal{A}$-子空间. x

命题 2 如果线性变换 $\mathcal{A}$ 与 $\mathcal{B}$ 可交换(即 $\mathcal{A}\mathcal{B}=\mathcal{B}\mathcal{A}$)，则 $\mathrm{Ker}\,\mathcal{B}$，$\mathrm{Im}\,\mathcal{B}$，$\mathcal{B}$ 的特征子空间都是 $\mathcal{A}$-子空间.

证明 任取 $\alpha\in\mathrm{Ker}\,\mathcal{B}$，则 $\mathcal{B}\alpha=0$.于是

$$\mathcal{B}(\mathcal{A}\alpha)=(\mathcal{B}\mathcal{A})\alpha=(\mathcal{A}\mathcal{B})\alpha=\mathcal{A}(\mathcal{B}\alpha)=\mathcal{A}(0)=0.$$

因此 $\mathcal{A}\alpha\in\mathrm{Ker}\,\mathcal{B}$，从而 $\mathrm{Ker}\,\mathcal{B}$ 是 $\mathcal{A}$-子空间.

类似地可证 $\mathrm{Im}\,\mathcal{B}$ 是 $\mathcal{A}$-子空间.

在 $\mathcal{B}$ 的特征子空间 V_{λ_i} 中任取一个向量 α，则 $\mathcal{B}\alpha=\lambda_i\alpha$.从而

$$\mathcal{B}(\mathcal{A}\alpha)=\mathcal{A}(\mathcal{B}\alpha)=\mathcal{A}(\lambda_i\alpha)=\lambda_i(\mathcal{A}\alpha),$$

因此 $\mathcal{A}\alpha\in V_{\lambda_i}$，从而 V_{λ_i} 是 $\mathcal{A}$-子空间. x

推论 3 设 $\mathcal{A}$ 是域 F 上线性空间 V 上的线性变换，则对于任意 $f(x)\in F[x]$，都有 $\mathrm{Ker}\,f(\mathcal{A})$，$\mathrm{Im}\,f(\mathcal{A})$，$f(\mathcal{A})$ 的特征子空间都是 $\mathcal{A}$-子空间. x

显然，线性变换 $\mathcal{A}$ 的不变子空间的交与和仍是 $\mathcal{A}$ 的不变子空间.

命题 4 设 $\mathcal{A}$ 是域 F 上线性空间 V 上的线性变换，$W=\langle\alpha_1,\alpha_2,\cdots,\alpha_m\rangle$ 是 V 的一个子空间，则 W 是 $\mathcal{A}$ 的不变子空间当且仅当 $\mathcal{A}\alpha_i\in W$，$i=1,2,\cdots,m$.

证明 由于 W 中任一向量 α 可表示成 $k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m$，因此，

W 是 $\mathcal{A}$-子空间

从 $\alpha\in W$ 可推出 $\mathcal{A}\alpha\in W$

$\mathcal{A}(k_1\alpha_1+k_2\alpha_2+\cdots+k_m\alpha_m)\in W$，" $k_i\in F$，$i=1,2,\cdots,m$

$\sum_{i=1}^{m}k_i\mathcal{A}\alpha_i\in W$，" $k_i\in F$，$i=1,2,\cdots,m$

$\mathcal{A}\alpha_i\in W$，$i=1,2,\cdots,m$ x

设 $\mathcal{A}$ 是域 F 上线性空间 V 上的线性变换，如果 W 是 $\mathcal{A}$-子空间，将 $\mathcal{A}$ 的

定义域限制到 W 上,且对应法则不变,则得到 W 上的一个线性变换,称它是 $\mathcal{A}$ 在 W 上的限制,记作 $\mathcal{A}|W$.由这个定义得

$$(\mathcal{A}|W)\alpha = \mathcal{A}\alpha, \quad \forall \alpha \in W. \tag{1}$$

如果 W 不是 $\mathcal{A}$-子空间,类似地可以有 $\mathcal{A}$ 在 W 上的限制,记作 $\mathcal{A}|W$,它不是 W 上的线性变换,而是 W 到 V 的一个线性映射.

设 $\mathcal{A}$ 是域 F 上有限维线性空间 V 上的线性变换,$\mathcal{A}$ 有没有非平凡的不变子空间与 $\mathcal{A}$ 的矩阵表示的形式有密切关系.

定理5　设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的线性变换,则 $\mathcal{A}$ 有非平凡的不变子空间当且仅当 V 中存在一个基,使得 $\mathcal{A}$ 在此基下的矩阵为分块上三角矩阵:

$$\begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix}. \tag{2}$$

证明　必要性.设 W 是 $\mathcal{A}$ 的非平凡不变子空间,在 W 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_r$,把它扩充成 V 的一个基:

$$\alpha_1,\alpha_2,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n,$$

则 $\mathcal{A}(\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n)$

$$=(\alpha_1,\cdots,\alpha_r,\alpha_{r+1},\cdots,\alpha_n)\begin{pmatrix} a_{11} & \cdots & a_{1r} & a_{1,r+1} & \cdots & a_{1n} \\ \cdots & & \cdots & \cdots & & \cdots \\ a_{r1} & \cdots & a_{rr} & a_{r,r+1} & \cdots & a_{rn} \\ 0 & \cdots & 0 & a_{r+1,r+1} & \cdots & a_{r+1,n} \\ \cdots & & \cdots & \cdots & & \cdots \\ 0 & \cdots & 0 & a_{n,r+1} & \cdots & a_{nn} \end{pmatrix}. \tag{3}$$

因此 $\mathcal{A}$ 在 V 的上述基下的矩阵为

$$\begin{pmatrix} A_1 & A_3 \\ 0 & A_2 \end{pmatrix},$$

其中 A_1 是 $\mathcal{A}|W$ 在 W 的一个基 $\alpha_1,\cdots,\alpha_r$ 下的矩阵.

充分性.设 $\mathcal{A}$ 在 V 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵是分块上三角矩阵(2),其中 A_1 是 r 级矩阵,$0<r<n$.令

$$W=\langle\alpha_1,\alpha_2,\cdots,\alpha_r\rangle. \tag{4}$$

则 $\mathcal{A}\alpha_i\in W$,$i=1,2,\cdots,r$.因此 W 是 $\mathcal{A}$-子空间,显然 W 是非平凡的.此时 A_1 是 $\mathcal{A}|W$ 在 W 的基 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 下的矩阵. ∎

定理6　设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的线性变换,则 V 能分解成 $\mathcal{A}$ 的一些非平凡不变子空间的直和当且仅当 V 中存在一个基,使得 $\mathcal{A}$ 在此基下的矩阵为分块对角矩阵:

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix}. \tag{5}$$

证明 必要性.设 V 是 $\mathcal{A}$ 的一些非平凡不变子空间的直和:

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_s. \tag{6}$$

在每个 $W_i(i=1,2,\cdots,s)$ 中取一个基 $\alpha_{i1},\cdots,\alpha_{ir_i}$,从(6)式得出,

$$\alpha_{11},\cdots,\alpha_{1r_1},\cdots,\alpha_{s1},\cdots,\alpha_{sr_s} \tag{7}$$

是 V 的一个基.由于 W_i 是 $\mathcal{A}$-子空间,因此

$$\mathcal{A}(\alpha_{i1},\cdots,\alpha_{ir_i}) = (\alpha_{i1},\cdots,\alpha_{ir_i})A_i,\quad i=1,2,\cdots,s. \tag{8}$$

从而 $\mathcal{A}$ 在(7)式给出的基下的矩阵为

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix}. \tag{9}$$

充分性.设 $\mathcal{A}$ 在 V 的一个基

$$\alpha_{11},\cdots,\alpha_{1r_1},\cdots,\alpha_{s1},\cdots,\alpha_{sr_s}$$

下的矩阵 $A=\mathrm{diag}\{A_1,A_2,\cdots,A_s\}$,其中 A_i 是 r_i 级方阵,$i=1,2,\cdots,s$.令 $W_i=\langle\alpha_{i1},\cdots,\alpha_{ir_i}\rangle$,$i=1,2,\cdots,s$.由于

$$\mathcal{A}(\alpha_{i1},\cdots,\alpha_{ir_i}) = (\alpha_{i1},\cdots,\alpha_{ir_i})A_i,\quad i=1,2,\cdots,s, \tag{10}$$

因此 $\mathcal{A}\alpha_{i1},\cdots,\mathcal{A}\alpha_{ir_i}\in W_i$.从而 W_i 是 $\mathcal{A}$-子空间,显然 W_i 是非平凡的,由于 W_i 的一个基 $\alpha_{i1},\cdots,\alpha_{ir_i}$ 当 $i=1,2,\cdots,s$ 时,合起来是 V 的一个基,因此 $W_1+W_2+\cdots+W_s$ 是直和,且 $V=W_1+W_2+\cdots+W_s$,从而 $V=W_1\oplus W_2\oplus\cdots\oplus W_s$. □

从定理 6 的证明中可看出,(5)式给出的矩阵中,A_i 就是 $\mathcal{A}|W_i$ 在 W_i 的一个基 $\alpha_{i1},\cdots,\alpha_{ir_i}$ 下的矩阵,其中 $i=1,2,\cdots,s$.

从定理 6 看出,如果 n 维线性空间 V 能分解成线性变换 $\mathcal{A}$ 的一些非平凡不变子空间的直和,那么 V 中存在一个基,使得 $\mathcal{A}$ 在此基下的矩阵为分块对角矩阵,这是比较简单的矩阵.问题是如何找出 $\mathcal{A}$ 的一些非平凡不变子空间?据推论 3,对于任意 $f(x)\in F[x]$,都有 $\mathrm{Ker}\,f(\mathcal{A})$ 是 $\mathcal{A}$ 的不变子空间.由此受到启发,能不能找到一些多项式 $f_1(x),f_2(x),\cdots,f_s(x)\in F[x]$,使得

$$V=\mathrm{Ker}\,f_1(\mathcal{A})\oplus\mathrm{Ker}\,f_2(\mathcal{A})\oplus\cdots\oplus\mathrm{Ker}\,f_s(\mathcal{A}). \tag{11}$$

为此我们首先来讨论,对于不同的一元多项式 $f_1(x)$, $f_2(x)$, x 用 A 代入得到的 $f_1(A)$与 $f_2(A)$的核之间有什么关系?

定理 7　设 A 是域 F 上线性空间 V 上的线性变换,在 $F[x]$中, $f(x) = f_1(x)f_2(x)$,且 $(f_1(x), f_2(x)) = 1$,则

$$\operatorname{Ker} f(A) = \operatorname{Ker} f_1(A) \oplus \operatorname{Ker} f_2(A). \tag{12}$$

证明　先证 $\operatorname{Ker} f(A) = \operatorname{Ker} f_1(A) + \operatorname{Ker} f_2(A)$.

因为 $f(x) = f_1(x)f_2(x)$,所以 $f(A) = f_1(A)f_2(A)$.任取 $\alpha_1 \in \operatorname{Ker} f_1(A)$,有 $f_1(A)\alpha_1 = 0$.从而有

$$f(A)\alpha_1 = (f_1(A)f_2(A))\alpha_1 = f_2(A)(f_1(A)\alpha_1) = 0,$$

因此 $\alpha_1 \in \operatorname{Ker} f(A)$.于是 $\operatorname{Ker} f_1(A) \subseteq \operatorname{Ker} f(A)$.同理可证, $\operatorname{Ker} f_2(A) \subseteq \operatorname{Ker} f(A)$.因此 $\operatorname{Ker} f_1(A) + \operatorname{Ker} f_2(A) \subseteq \operatorname{Ker} f(A)$.

由于 $(f_1(x), f_2(x)) = 1$,因此存在 $u_1(x), u_2(x) \in F[x]$,使得 $u_1(x)f_1(x) + u_2(x)f_2(x) = 1$. x 用 A 代入,得

$$u_1(A)f_1(A) + u_2(A)f_2(A) = I. \tag{13}$$

任取 $\alpha \in \operatorname{Ker} f(A)$,则由(13)式得

$$\alpha = I\alpha = u_1(A)f_1(A)\alpha + u_2(A)f_2(A)\alpha. \tag{14}$$

令　$\alpha_1 = u_2(A)f_2(A)\alpha$, $\alpha_2 = u_1(A)f_1(A)\alpha$.

由于

$$f_1(A)\alpha_1 = f_1(A)u_2(A)f_2(A)\alpha = u_2(A)f(A)\alpha = 0,$$

因此, $\alpha_1 \in \operatorname{Ker} f_1(A)$,同理可证, $\alpha_2 \in \operatorname{Ker} f_2(A)$.于是从(14)式得

$$\alpha = \alpha_1 + \alpha_2 \in \operatorname{Ker} f_1(A) + \operatorname{Ker} f_2(A),$$

所以

$$\operatorname{Ker} f(A) = \operatorname{Ker} f_1(A) + \operatorname{Ker} f_2(A).$$

再证　$\operatorname{Ker} f_1(A) \cap \operatorname{Ker} f_2(A) = 0$.任取 $\beta \in \operatorname{Ker} f_1(A) \cap \operatorname{Ker} f_2(A)$,则

$$\beta = I\beta = u_1(A)f_1(A)\beta + u_2(A)f_2(A)\beta = 0 + 0 = 0.$$

因此　$\operatorname{Ker} f_1(A) \cap \operatorname{Ker} f_2(A) = 0$.综上所述,得

$$\operatorname{Ker} f(A) = \operatorname{Ker} f_1(A) \oplus \operatorname{Ker} f_2(A).$$

□

用数学归纳法可以把定理 7 推广成:

推论 8　设 A 是域 F 上线性空间 V 上的线性变换,在 $F[x]$中, $f(x) = f_1(x)f_2(x)\cdots f_s(x)$,其中 $f_1(x), f_2(x), \cdots, f_s(x)$两两互素,则

$$\operatorname{Ker} f(A) = \operatorname{Ker} f_1(A) \oplus \operatorname{Ker} f_2(A) \oplus \cdots \oplus \operatorname{Ker} f_s(A). \tag{15}$$

□

由于 $\operatorname{Ker} 0 = V$,因此从推论 8 受到启发,如果能找到一个多项式 $f(x)$,使得 $f(A) = 0$,那么空间 V 就能分解成

$$V = \mathrm{Ker} f_1(A) \oplus \cdots \oplus \mathrm{Ker} f_s(A), \tag{16}$$

其中，$f_1(x)\cdots f_s(x) = f(x)$，并且 $f_1(x),\cdots,f_s(x)$两两互素.

定义 2　设 A 是 V 上的线性变换，如果 F 上的一元多项式 $f(x)$，使得 $f(A)=0$，则称 $f(x)$是 A 的一个零化多项式.

换句话说，如果 $f(x)$在 $F[A]$中的一个根是 A，则称 $f(x)$是 A 的一个零化多项式.

如果 V 是有限维的，那么 V 上的任意一个线性变换 A 都有非零的零化多项式.理由如下：设 $\dim V = n$，则 $\dim \mathrm{Hom}(V, V) = n^2$.从而 $I, A, A^2, \cdots, A^{n^2}$ 一定线性相关.于是有 F 中不全为零的元素 $k_0, k_1, \cdots, k_{n^2}$，使得

$$k_0 I + k_1 A + k_2 A^2 + \cdots + k_{n^2} A^{n^2} = 0. \tag{17}$$

令
$$f(x) = k_0 + k_1 x + k_2 x^2 + \cdots + k_{n^2} x^{n^2}$$

由(17)式得，$f(A)=0$.即 $f(x)$是 A 的一个非零的零化多项式.

定义 3　设 A 是域 F 上一个 n 级矩阵，如果 $F[x]$中的一个多项式 $f(x)$使得 $f(A)=0$，则称 $f(x)$是矩阵 A 的一个零化多项式.

设 A 是域 F 上 n 维线性空间 V 上的线性变换，且 A 在 V 的一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵为 A.则

$$k_0 I + k_1 A + \cdots + k_m A^m = 0 \qquad k_0 I + k_1 A + \cdots + k_m A^m = 0,$$

其中 $k_0, k_1, \cdots, k_m \in F$，$m$ 是非负整数.从而

$$f(A) = 0 \qquad f(A) = 0,$$

即　　$f(x)$是 A 的零化多项式　　$f(x)$是 A 的零化多项式.

习　题　9.5

1. 在 K^4 中，设线性变换 A 在基 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 下的矩阵是

$$A = \begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & 4 & -2 \\ 2 & -1 & 0 & 1 \\ 2 & -1 & -1 & 2 \end{pmatrix},$$

令 $W = \langle \alpha_1 + 2\alpha_2, \alpha_2 + \alpha_3 + 2\alpha_4 \rangle$.证明：$W$ 是 A - 子空间.

2. 设 W 是 V 上可逆线性变换 A 的有限维不变子空间，证明：

(1) $A|W$ 是 W 上的可逆线性变换；

(2) W 也是 A^{-1}的不变子空间，并且$(A|W)^{-1} = A^{-1}|W$.

3. 设 V 是复数域上的 n 维线性空间，A, B 是 V 上的线性变换，并且 $AB = BA$.证明：A

与 $\mathcal{B}$ 至少有一个公共的特征向量.

4. 设 V 是域 F 上的 n 维线性空间,线性变换 $\mathcal{A}$ 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵为

$$A = \begin{pmatrix} a & 1 & 0 & \cdots & 0 & 0 \\ 0 & a & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & a & 1 \\ 0 & 0 & 0 & \cdots & 0 & a \end{pmatrix},$$

(1) 证明:如果 α_n 属于 $\mathcal{A}$ 的不变子空间 W,则 $W = V$;

(2) 证明:α_1 属于 $\mathcal{A}$ 的任意一个非零不变子空间;

(3) 证明:V 不能分解成 $\mathcal{A}$ 的两个非平凡不变子空间的直和;

*(4) 求 $\mathcal{A}$ 的所有不变子空间.

5. 设 V 是实数域上 n 维线性空间,证明:V 的任一线性变换 $\mathcal{A}$ 必有一个 1 维不变子空间或者 2 维不变子空间.

6. 设 V 是平面上以定点 O 为起点的所有向量组成的实数域上的 2 维线性空间,$\mathcal{A}$ 是绕 O 点转角为 θ 的旋转,其中 $\theta \neq k\pi, k \in \mathbf{Z}$.证明:$\mathcal{A}$ 没有非平凡的不变子空间.

*7. 设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的线性变换,λ_0 是 $\mathcal{A}$ 的一个特征值.λ_0 作为 $\mathcal{A}$ 的特征多项式的根的重数称为 λ_0 的代数重数;$\mathcal{A}$ 的属于 λ_0 的特征子空间 V_{λ_0} 的维数称为 λ_0 的几何重数.证明:λ_0 的几何重数不超过它的代数重数.

*8. 设 $\mathcal{A}$ 是复数域上 n 维线性空间 V 上的线性变换.证明:$\mathcal{A}$ 可对角化的充分必要条件是,$\mathcal{A}$ 的每一个特征值的几何重数等于它的代数重数.

9. 设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的线性变换,证明:$\mathcal{A}$ 是幂等变换的充分必要条件是

$$\mathrm{rank}(\mathcal{A}) + \mathrm{rank}(\mathcal{A} - \mathcal{I}) = n.$$

*10. 设 A 是域 F 上的 n 级矩阵.证明:A 是对合矩阵(即 $A^2 = I$)的充分必要条件是

$$\mathrm{rank}(A + I) + \mathrm{rank}(A - I) = n.$$

11. 写出数乘变换 k 的一个非零的零化多项式.

12. 证明:如果线性变换 $\mathcal{A}$ 有一个 1 次的零化多项式,则 $\mathcal{A}$ 一定是数乘变换.

13. 求下述 2 级矩阵 A 的一个非零的零化多项式:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

14. 求下述 n 级矩阵 A 的一个非零的零化多项式:

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

*15. 设 W 是 V 上线性变换 $\mathcal{A}$ 的非零不变子空间,令

$$\overline{A}: V/W \longrightarrow V/W$$
$$\alpha + W \longmapsto A\alpha + W,$$

证明: $\overline{A}$ 是商空间 V/W 上的一个线性变换.

§6 Hamilton - Cayley 定理

设数域 K 上的 2 级矩阵 A 为

$$A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix},$$

则

$$A^2 = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.$$

即 $A^2 - I = 0$. 因此 $f(\lambda) = \lambda^2 - 1$ 是 A 的一个零化多项式. 由于

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & -2 \\ 0 & \lambda + 1 \end{vmatrix} = (\lambda - 1)(\lambda + 1) = \lambda^2 - 1.$$

因此 A 的特征多项式 $f(\lambda) = \lambda^2 - 1$ 就是 A 的一个零化多项式. 这个结论对于域 F 上任意一个 n 级矩阵 A 成立吗? 为了讨论这个问题, 我们需要把域上的矩阵的概念推广到整环上的矩阵.

设 R 是一个整环, R 中 n^2 个元素排成的 n 行 n 列的一张表称为 R 上的一个 n 级矩阵. 类似于域 F 上的矩阵的加法、纯量乘法、乘法的定义, 我们可以定义整环 R 上的矩阵的加法、纯量乘法、乘法, 而且这些运算满足与域 F 上矩阵一样的运算法则. 与域 F 上 n 级矩阵的行列式的定义类似, 可以定义整环 R 上的 n 级矩阵的行列式, 而且本套教材上册第 2 章讲的行列式的性质、行列式按一行(列)展开的定理, 对于整环 R 上的 n 级矩阵的行列式也成立, 从而对于整环 R 上的矩阵 A 也有

$$AA^* = A^*A = |A|I,$$

其中 A^* 是 A 的伴随矩阵.

域 F 上的一元多项式环 $F[\lambda]$ 是一个整环, 因此可以考虑环 $F[\lambda]$ 上的 n 级矩阵, 称这样的矩阵为 λ - 矩阵, 记成 $A(\lambda)$, $B(\lambda)$, $\cdots$, 等. 例如

$$A(\lambda) = \begin{pmatrix} 2\lambda^3 + \lambda^2 + 1 & \lambda^2 - 3 \\ \lambda^3 - 1 & 2\lambda + 5 \end{pmatrix} \tag{1}$$

是一个 λ - 矩阵, 按照整环上矩阵的加法与纯量乘法, 可以把(1)中的 λ - 矩阵 $A(\lambda)$ 写成

$$A(\lambda) = \begin{pmatrix} 2\lambda^3 & 0 \\ \lambda^3 & 0 \end{pmatrix} + \begin{pmatrix} \lambda^2 & \lambda^2 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 2\lambda \end{pmatrix} + \begin{pmatrix} 1 & -3 \\ -1 & 5 \end{pmatrix}$$

$$= \lambda^3 \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix} + \lambda^2 \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + \lambda \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 1 & -3 \\ -1 & 5 \end{pmatrix}, \tag{2}$$

其中 λ^i 的“系数矩阵”是域 F 上的矩阵.

当我们把两个 λ-矩阵 $A(\lambda)$ 和 $B(\lambda)$ 都写成(2)那样的形式时,根据两个一元多项式相等的定义以及两个 λ-矩阵相等的定义,可以推出,$A(\lambda)$ 与 $B(\lambda)$ 相等当且仅当它们的 λ^i 的“系数矩阵”对应相等,$i = 0, 1, 2, \cdots$.

有了上述准备知识后,我们来证明下面的 Hamilton - Cayloy 定理.

Hamilton - Cayley 定理 设 A 是域 F 上的 n 级矩阵,则 A 的特征多项式 $f(\lambda)$ 是 A 的一个零化多项式.

证明 设 $B(\lambda)$ 是 λ-矩阵$(\lambda I - A)$的伴随矩阵,则

$$B(\lambda)(\lambda I - A) = |\lambda I - A| I = f(\lambda) I. \tag{3}$$

因为 $B(\lambda)$ 的元素是矩阵 $\lambda I - A$ 的元素(一次或零次多项式)的代数余子式,所以 $B(\lambda)$ 的元素都是次数不超过 $n-1$ 的一元多项式.因此 $B(\lambda)$ 可以写成

$$B(\lambda) = \lambda^{n-1} B_{n-1} + \lambda^{n-2} B_{n-2} + \cdots + \lambda B_1 + B_0, \tag{4}$$

其中 $B_{n-1}, B_{n-2}, \cdots, B_1, B_0$ 都是域 F 上的 n 级矩阵.

直接计算得

$$\begin{aligned} & B(\lambda)(\lambda I - A) \\ &= (\lambda^{n-1} B_{n-1} + \lambda^{n-2} B_{n-2} + \cdots + \lambda B_1 + B_0)(\lambda I - A) \\ &= \lambda^n B_{n-1} + \lambda^{n-1}(B_{n-2} - B_{n-1} A) + \cdots + \lambda(B_0 - B_1 A) - B_0 A. \end{aligned} \tag{5}$$

由于 A 的特征多项式 $f(\lambda)$ 的次数是 n,于是可以设

$$f(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0. \tag{6}$$

从而

$$f(\lambda) I = \lambda^n I + a_{n-1}\lambda^{n-1} I + \cdots + a_1 \lambda I + a_0 I. \tag{7}$$

据(3),(5),(7)式得

$$\begin{aligned} & B_{n-1} = I, \\ & B_{n-2} - B_{n-1} A = a_{n-1} I, \\ & \cdots\cdots\cdots\cdots \\ & B_0 - B_1 A = a_1 I, \\ & -B_0 A = a_0 I. \end{aligned} \tag{8}$$

用 $A^n, A^{n-1}, \cdots, A, I$ 依次从右边乘(8)的第 1 式,第 2 式,…,第 n 式,第 $n+1$ 式,得

$$\begin{aligned} & B_{n-1} A^n = A^n, \\ & B_{n-2} A^{n-1} - B_{n-1} A^n = a_{n-1} A^{n-1}, \\ & \cdots\cdots\cdots\cdots \\ & B_0 A - B_1 A^2 = a_1 A, \\ & -B_0 A = a_0 I, \end{aligned} \tag{9}$$

把(9)的 $n+1$ 个式子一起加起来,左边为零,右边即为 $f(\mathcal{A})$.因此 $f(\mathcal{A})=0$.　x

用线性变换的语言叙述上述定理,即

Hamilton - Cayley 定理　设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的一个线性变换,则 $\mathcal{A}$ 的特征多项式 $f(\lambda)$ 是 $\mathcal{A}$ 的一个零化多项式.　x

习　题　9.6

1. 设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的线性变换.

证明:如果 $\mathcal{A}$ 的特征多项式 $f(\lambda)$ 在 $F[\lambda]$ 中可以分解成

$$f(\lambda)=(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_s)^{r_s},$$

则
$$V=\mathrm{Ker}(\mathcal{A}-\lambda_1\mathcal{I})^{r_1}\oplus\mathrm{Ker}(\mathcal{A}-\lambda_2\mathcal{I})^{r_2}\oplus\cdots\oplus\mathrm{Ker}(\mathcal{A}-\lambda_s\mathcal{I})^{r_s},$$

其中 $\mathrm{Ker}(\mathcal{A}-\lambda_j\mathcal{I})^{r_j}$ 称为 $\mathcal{A}$ 的根子空间,$j=1,2,\cdots,s$.

2. 证明:对于域 F 上的 n 级可逆矩阵 A,存在 F 中元素 $k_0,k_1,\cdots,k_{n-1}$,使得

$$A^{-1}=k_{n-1}A^{n-1}+\cdots+k_1A+k_0I.$$

3. 求下列复矩阵的一个非零的零化多项式:

(1) $\begin{pmatrix}1 & 2\\ -1 & 4\end{pmatrix}$;　(2) $\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}$.

*4. 设 A,B 分别是 n 级、m 级复矩阵,证明:矩阵方程 $AX-XB=0$ 只有零解的充分必要条件是,A 与 B 没有公共的特征值.

§7　线性变换的最小多项式

从 Hamilton - Cayley 定理知道,有限维线性空间 V 上的线性变换 $\mathcal{A}$ 的特征多项式 $f(\lambda)$ 是 $\mathcal{A}$ 的一个零化多项式.这一节我们来讨论 $\mathcal{A}$ 有没有别的零化多项式?如果还有的话,那么 $\mathcal{A}$ 的零化多项式之间有什么关系?

定义 1　设 $\mathcal{A}$ 是域 F 上线性空间 V 的一个线性变换,在 $\mathcal{A}$ 的所有非零的零化多项式中,次数最低的首项系数为 1 的多项式称为 $\mathcal{A}$ 的最小多项式.

命题 1　线性空间 V 上的线性变换 $\mathcal{A}$ 的最小多项式是唯一的.

证明　设 $m_1(\lambda)$ 和 $m_2(\lambda)$ 都是 $\mathcal{A}$ 的最小多项式,则它们的次数相同并且首项系数都为 1,从而 $h(\lambda)=m_1(\lambda)-m_2(\lambda)$ 的次数比 $m_1(\lambda)$ 的低.由于 $h(\mathcal{A})=m_1(\mathcal{A})-m_2(\mathcal{A})=0$,所以 $h(\lambda)$ 也是 $\mathcal{A}$ 的一个零化多项式.据最小多项式的定义得,$h(\lambda)=0$.即

$$m_1(\lambda)=m_2(\lambda).$$　x

命题2　线性空间 V 上的线性变换 $\mathcal{A}$ 的任一零化多项式 $g(\lambda)$ 是 $\mathcal{A}$ 的最小多项式 $m(\lambda)$ 的倍式.

证明　作带余除法,得

$$g(\lambda)=h(\lambda)m(\lambda)+r(\lambda),\quad \deg r(\lambda)<\deg m(\lambda).$$

不定元 λ 用 $\mathcal{A}$ 代入,从上式得

$$g(\mathcal{A})=h(\mathcal{A})m(\mathcal{A})+r(\mathcal{A}).$$

由于 $g(\mathcal{A})=0$, $m(\mathcal{A})=0$,所以 $r(\mathcal{A})=0$.据最小多项式的定义得,$r(\lambda)=0$.因此 $g(\lambda)$ 是 $m(\lambda)$ 的倍式. x

命题3　设 $\mathcal{A}$ 是域 F 上有限维线性空间 V 上的线性变换,则 $\mathcal{A}$ 的最小多项式 $m(\lambda)$ 与特征多项式 $f(\lambda)$ 在 F 中有相同的根(重数可以不同).

证明　据 Hamilton - Cayley 定理和性质2得,$f(\lambda)=h(\lambda)m(\lambda)$,对于某个 $h(\lambda)\in F[\lambda]$.从而 $m(\lambda)$ 的每一个根都是 $f(\lambda)$ 的根.

反之,设 λ_0 是 $f(\lambda)$ 在 F 中的一个根,则 λ_0 是 $\mathcal{A}$ 的一个特征值.从而存在 $\xi\in V$ 且 $\xi\neq 0$,使得 $\mathcal{A}\xi=\lambda_0\xi$.设

$$m(\lambda)=c_0+c_1\lambda+\cdots+c_s\lambda^s,$$

则

$$0=m(\mathcal{A})\xi=(c_0\mathcal{I}+c_1\mathcal{A}+\cdots+c_s\mathcal{A}^s)\xi$$
$$=c_0\xi+c_1\lambda_0\xi+\cdots+c_s\lambda_0^s\xi=m(\lambda_0)\xi.$$

由此得出,$m(\lambda_0)=0$,即 λ_0 是 $m(\lambda)$ 的一个根. x

类似地,可以定义域 F 上 n 级矩阵 A 的最小多项式.

设 n 维线性空间 V 上的线性变换 $\mathcal{A}$ 在 V 的一个基下的矩阵是 A,在§5中我们已指出,$g(\lambda)$ 是 $\mathcal{A}$ 的零化多项式当且仅当 $g(\lambda)$ 是 A 的零化多项式.由此得出,$m(\lambda)$ 是 $\mathcal{A}$ 的最小多项式当且仅当 $m(\lambda)$ 是 A 的最小多项式.于是从性质3得出,

推论4　域 F 上 n 级矩阵 A 的最小多项式 $m(\lambda)$ 与特征多项式 $f(x)$ 在 F 中有相同的根(重数可以不同). x

由于相似的矩阵可以看作是同一个线性变换 $\mathcal{A}$ 在 V 的不同基下的矩阵,因此有下述结论:

推论5　相似的矩阵有相同的最小多项式. x

如何求线性变换 $\mathcal{A}$(或 n 级矩阵 A)的最小多项式?一种方法是:先找出 $\mathcal{A}$ 的一个零化多项式 $g(\lambda)$,然后分析 $g(\lambda)$ 的哪个因式是 $\mathcal{A}$ 的最小多项式.

定义2　设 $\mathcal{A}$ 是域 F 上线性空间 V 上的线性变换,如果存在一个正整数 l,使得 $\mathcal{A}^l=0$,则称 $\mathcal{A}$ 是幂零变换;使得 $\mathcal{A}^l=0$ 成立的最小正整数 l 称为 $\mathcal{A}$ 的幂零指数.

例1　求幂零指数为 l 的幂零变换 $\mathcal{A}$ 的最小多项式.

解　由于 $\mathcal{A}^l=0$，因此 $g(\lambda)=\lambda^l$ 是幂零变换 $\mathcal{A}$ 的一个零化多项式．据幂零指数的定义知，当 $0<s<l$ 时，$\mathcal{A}^s\neq 0$．从而 λ^l 是 $\mathcal{A}$ 的最小多项式．

例 2　设域 F 上的 r 级矩阵 A 为

$$A=\begin{pmatrix}0&1&0&\cdots&0&0\\0&0&1&\cdots&0&0\\\cdots&\cdots&\cdots&&\cdots&\cdots\\0&0&0&\cdots&0&1\\0&0&0&\cdots&0&0\end{pmatrix},$$

求 A 的最小多项式．

解　设 W 是域 F 上 r 维线性空间，V 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_r$，建立 W 上的一个线性变换 $\mathcal{A}$，使得 $\mathcal{A}(\alpha_1,\alpha_2,\cdots,\alpha_r)=(\alpha_1,\alpha_2,\cdots,\alpha_r)A$．

则
$$\mathcal{A}\alpha_1=0,\ \mathcal{A}\alpha_2=\alpha_1,\ \mathcal{A}\alpha_3=\alpha_2,\cdots,\ \mathcal{A}\alpha_r=\alpha_{r-1}.$$

从而
$$\mathcal{A}^2\alpha_2=\mathcal{A}(\mathcal{A}\alpha_2)=\mathcal{A}\alpha_1=0,$$
$$\mathcal{A}^3\alpha_3=\mathcal{A}^2(\mathcal{A}\alpha_3)=\mathcal{A}^2(\alpha_2)=0,$$
$$\cdots\cdots\cdots\cdots$$
$$\mathcal{A}^r\alpha_r=\mathcal{A}^{r-1}(\mathcal{A}\alpha_r)=\mathcal{A}^{r-1}(\alpha_{r-1})=0.$$

因此 $\mathcal{A}$ 是幂零变换，且幂零指数为 r．从例 1 知道，$\mathcal{A}$ 的最小多项式是 λ^r．从而矩阵 A 的最小多项式是 λ^r．

从例 2 的求解过程看出，例 2 的矩阵 A 是幂零矩阵，且幂零指数等于这个矩阵的级数，也可以通过直接计算得出，$A^r=0$；当 $0<s<r$ 时，$A^s\neq 0$．

比例 2 的矩阵更一般地考虑下述矩阵：

定义 3　域 F 上的一个 r 级矩阵如果形如

$$\begin{pmatrix}a&1&0&\cdots&0&0\\0&a&1&\cdots&0&0\\\cdots&\cdots&\cdots&&\cdots&\cdots\\0&0&0&\cdots&a&1\\0&0&0&\cdots&0&a\end{pmatrix},\tag{1}$$

则称它为一个 r 级 Jordan 块，记作 $J_r(a)$，其中 a 是主对角线上的元素．

例 2 的矩阵是主对角元为 0 的 r 级 Jordan 块 $J_r(0)$．

1 级 Jordan 块 $J_1(a)$ 就是 1 级矩阵 (a)．

例 3　求 r 级 Jordan 块 $J_r(a)$ 的最小多项式．

解　先求 $J_r(a)$ 的特征多项式 $f(\lambda)$：

$$f(\lambda)=|\lambda I-J_r(a)|=\begin{vmatrix}\lambda-a & -1 & 0 & \cdots & 0 & 0\\ 0 & \lambda-a & -1 & \cdots & 0 & 0\\ \cdots & \cdots & \cdots & & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & \lambda-a & -1\\ 0 & 0 & 0 & \cdots & 0 & \lambda-a\end{vmatrix}=(\lambda-a)^r.$$

由于 $J_r(a)$的最小多项式 $m(\lambda)$是 $f(\lambda)$的因式,因此

$$m(\lambda)=(\lambda-a)^s,\text{对于某个 } s(0<s\leq r).$$

λ用 $J_r(a)$代入,得

$$0=m(J_r(a))=(J_r(a)-aI)^s=(J_r(0))^s.$$

由于 $J_r(0)$的幂零指数为 r,因此 $s=r$.从而 $J_r(a)$的最小多项式 $m(\lambda)=(\lambda-a)^r$.

下面的定理 6 给出了求线性变换 $\mathcal{A}$ 的最小多项式的又一种方法.

定理 6　设 $\mathcal{A}$ 是域 F 上线性空间 V 上的线性变换.如果 V 能分解成 $\mathcal{A}$ 的一些非平凡不变子空间的直和:

$$V=W_1\oplus W_2\oplus\cdots\oplus W_s,\tag{2}$$

则 $\mathcal{A}$ 的最小多项式 $m(\lambda)$为

$$m(\lambda)=[m_1(\lambda),m_2(\lambda),\cdots,m_s(\lambda)],\tag{3}$$

其中 $m_j(\lambda)$是 W_j 上的线性变换 $\mathcal{A}|W_j$ 的最小多项式,$j=1,2,\cdots,s$. $[m_1(\lambda),m_2(\lambda),\cdots,m_s(\lambda)]$ 是 $m_1(\lambda),m_2(\lambda),\cdots,m_s(\lambda)$的最小公倍式.

证明　记 $g(\lambda)=[m_1(\lambda),m_2(\lambda),\cdots,m_s(\lambda)]$,则

$$g(\lambda)=h_j(\lambda)m_j(\lambda),\quad j=1,2,\cdots,s.$$

先证 $g(\lambda)|m(\lambda)$,为此只要证 $m_j(\lambda)|m(\lambda)$, $j=1,2,\cdots,s$.任取 $\alpha_j\in W_j$,有

$$m(\mathcal{A}|W_j)\alpha_j=m(\mathcal{A})\alpha_j=0\alpha_j=0,$$

因此 $m(\mathcal{A}|W_j)=0$,即 $m(\lambda)$是 W_j上线性变换 $\mathcal{A}|W_j$的一个零化多项式.从而 $m_j(\lambda)|m(\lambda)$.由于 $j=1,2,\cdots,s$,因此 $m(\lambda)$是 $m_1(\lambda),m_2(\lambda),\cdots,m_s(\lambda)$的公倍式,从而它们的最小公倍式 $g(\lambda)|m(\lambda)$.

再证 $m(\lambda)|g(\lambda)$,为此只要证 $g(\lambda)$是 $\mathcal{A}$ 的零化多项式.任取 $\alpha\in V$.由于 $V=W_1\oplus W_2\oplus\cdots\oplus W_s$,因此

$$\alpha=\alpha_1+\alpha_2+\cdots+\alpha_s,\quad\alpha_i\in W_i,\ i=1,2,\cdots,s.$$

从而

$$g(\mathcal{A})\alpha=g(\mathcal{A})\sum_{j=1}^{s}\alpha_j=\sum_{j=1}^{s}g(\mathcal{A})\alpha_j=\sum_{j=1}^{s}h_j(\mathcal{A})m_j(\mathcal{A})\alpha_j$$

$$= \sum_{j=1}^{s} h_j(\mathcal{A})\, m_j(\mathcal{A}|W_j)\alpha_j = \sum_{j=1}^{s} h_j(\mathcal{A})[0\alpha_j] = 0.$$

因此 $g(\mathcal{A})=0$.从而 $g(\lambda)$是 $\mathcal{A}$ 的零化多项式.于是 $m(\lambda)\,|\,g(\lambda)$.

综上所述得, $m(\lambda) = g(\lambda) = [m_1(\lambda), m_2(\lambda), \cdots, m_s(\lambda)]$. x

推论 7　设 A 是域 F 上一个 n 级分块对角矩阵,即 $A = \mathrm{diag}\{A_1, A_2, \cdots, A_s\}$,设 A_j的最小多项式是 $m_j(\lambda)$,则 A 的最小多项式 $m(\lambda)$为

$$m(\lambda) = [m_1(\lambda), m_2(\lambda), \cdots, m_s(\lambda)].$$

证明　设 V 是域 F 上 n 维线性空间, V 中取一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$,建立 V 上的一个线性变换 $\mathcal{A}$,使得

$$\mathcal{A}(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)A.$$

由于 A 是分块对角矩阵,据§5 的定理 6 得, V 能分解成 $\mathcal{A}$ 的一些非平凡不变子空间的直和:

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_s,$$

并且 A_j就是 $\mathcal{A}|W_j$在 W_j的一个基下的矩阵.从而 $\mathcal{A}|W_j$的最小多项式就是 A_j的最小多项式 $m_j(\lambda)$, $j=1,2,\cdots,s$.又 A 的最小多项式就是 $\mathcal{A}$ 的最小多项式 $m(\lambda)$.据定理 6 得

$$m(\lambda) = [m_1(\lambda), m_2(\lambda), \cdots, m_s(\lambda)].$$ x

定义 4　由若干个 Jordan 块组成的分块对角矩阵称为 Jordan 形矩阵.

对角矩阵可以看成是由 1 级 Jordan 块组成的 Jordan 形矩阵.

例 4　设 A 是 Jordan 形矩阵:

$$A = \mathrm{diag}\{J_3(a), J_5(a), J_2(b)\},$$

求 A 的最小多项式.

解　$J_3(a)$, $J_5(a)$的最小多项式分别是$(\lambda-a)^3$, $(\lambda-a)^5$, $J_2(b)$的最小多项式是$(\lambda-b)^2$.因此 A 的最小多项式为

$$m(\lambda) = [(\lambda-a)^3, (\lambda-a)^5, (\lambda-b)^2] = (\lambda-a)^5(\lambda-b)^2.$$

线性变换的最小多项式在研究线性变换的结构中起着十分重要的作用.现在先利用最小多项式给出线性变换可对角化的一个充分必要条件.然后用最小多项式研究不可以对角化的线性变换的结构.

定理 8　设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的线性变换,则 $\mathcal{A}$ 可对角化的充分必要条件是, $\mathcal{A}$ 的最小多项式 $m(\lambda)$在 $F[\lambda]$中能分解成不同的一次因式的乘积.

证明　必要性.设 $\mathcal{A}$ 可对角化,则

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_s},$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是 A 的所有不同的特征值. 任取 $\alpha_j \in V_{\lambda_j}$, 有 $A\alpha_j = \lambda_j \alpha_j$, 从而 $(A - \lambda_j I)\alpha_j = 0$, 于是 $A|V_{\lambda_j} - \lambda_j I \ \alpha_j = 0$. 因此 $A|V_{\lambda_j} - \lambda_j I = 0$, 从而 $\lambda - \lambda_j$ 是 $A|V_{\lambda_j}$ 的一个零化多项式, 由于 $A|V_{\lambda_j}$ 的最小多项式 $m_j(\lambda)$ 是 $\lambda - \lambda_j$ 的因式, 因此 $m_j(\lambda) = \lambda - \lambda_j$. 从而 A 的最小多项式 $m(\lambda)$ 为

$$\begin{aligned} m(\lambda) &= \ \lambda - \lambda_1, \lambda - \lambda_2, \cdots, \lambda - \lambda_s \\ &= (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_s). \end{aligned}$$

充分性, 设 A 的最小多项式 $m(\lambda)$ 为

$$m(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_s), \tag{4}$$

其中 $\lambda_1, \lambda_2, \cdots, \lambda_s$ 是 F 中两两不同的元素. 据 §5 的推论 8 得,

$$\mathrm{Ker}\ m(A) = \mathrm{Ker}(A - \lambda_1 I) \oplus \mathrm{Ker}(A - \lambda_2 I) \oplus \cdots \oplus \mathrm{Ker}(A - \lambda_s I).$$

显然, $\mathrm{Ker}\ m(A) = \mathrm{Ker}\, 0 = V$. 由于

$$\alpha \in \mathrm{Ker}(A - \lambda_j I) \qquad (A - \lambda_j I)\alpha = 0 \qquad \alpha \in V_{\lambda_j},$$

因此 $\mathrm{Ker}(A - \lambda_j I) = V_{\lambda_j}$, $j = 1, 2, \cdots, s$. 从而

$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_s},$$

因此 A 可对角化. x

推论 9　域 F 上 n 级矩阵 A 可对角化的充分必要条件是, A 的最小多项式在 $F[\lambda]$ 中能分解成不同的一次因式的乘积. x

例 5　证明: 幂等矩阵一定可对角化.

证明　设 A 是幂等矩阵, 则 $A^2 = A$, 从而 $A^2 - A = 0$. 因此 $\lambda^2 - \lambda$ 是 A 的一个零化多项式, 显然

$$\lambda^2 - \lambda = \lambda(\lambda - 1).$$

由于 A 的最小多项式 $m(\lambda)$ 是 $\lambda^2 - \lambda$ 的因式, 因此 $m(\lambda)$ 也可以分解成不同的一次因式的乘积, 从而 A 可对角化. x

例 6　证明: 幂零指数 $l > 1$ 的幂零矩阵一定不可以对角化.

证明　设 A 是幂零指数 $l > 1$ 的幂零矩阵, 则 A 的最小多项式 $m(\lambda) = \lambda^l$. 由于 $l > 1$, 因此 A 不可对角化. x

从例 5 和例 6 看到, 在一些情形下, 利用最小多项式来判别线性变换或矩阵是否可对角化, 往往很简洁.

利用最小多项式还可以研究不可对角化的线性变换的结构, 即研究能不能找到 n 维线性空间 V 的一个基, 使得 V 上的线性变换 A 在此基下的矩阵 A 最简单?

设 A 是域 F 上 n 维线性空间 V 上的线性变换, 如果 A 的最小多项式 $m(\lambda)$ 在 $F[\lambda]$ 中能分解成一次因式的乘积:

$$m(\lambda)=(\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\cdots(\lambda-\lambda_s)^{l_s}, \tag{5}$$

其中 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是 F 中两两不同的元素.则

$$\operatorname{Ker} m(\mathcal{A})=\operatorname{Ker}(\mathcal{A}-\lambda_1\mathcal{I})^{l_1}\oplus\operatorname{Ker}(\mathcal{A}-\lambda_2\mathcal{I})^{l_2}\oplus\cdots\oplus\operatorname{Ker}(\mathcal{A}-\lambda_s\mathcal{I})^{l_s}. \tag{6}$$

显然,$\operatorname{Ker} m(\mathcal{A})=\operatorname{Ker} 0=V$.记 $W_j=\operatorname{Ker}(\mathcal{A}-\lambda_j\mathcal{I})^{l_j}$,据§5 的推论 3 得,$W_j$ 是 $\mathcal{A}$ 的不变子空间,$j=1,2,\cdots,s$.由(6)式得,

$$V=W_1\oplus W_2\oplus\cdots\oplus W_s. \tag{7}$$

在 W_j 中取一个基,$j=1,2,\cdots,s$,它们合起来成为 V 的一个基,$\mathcal{A}$ 在 V 的这个基下的矩阵 A 为分块对角矩阵:

$$A=\begin{pmatrix}A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s\end{pmatrix},$$

其中 A_j 是 $\mathcal{A}|W_j$ 在 W_j 的上述基下的矩阵.为了使矩阵 A 最简单,就应当使 A_j 最简单,$j=1,2,\cdots,s$.为此我们来研究 W_j 上的线性变换 $\mathcal{A}|W_j$.我们断言 $\mathcal{A}|W_j$ 的最小多项式是$(\lambda-\lambda_j)^{l_j}$,理由如下:

任取 $\alpha_j\in W_j$,由于 $W_j=\operatorname{Ker}(\mathcal{A}-\lambda_j\mathcal{I})^{l_j}$,因此

$$(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{l_j}\alpha_j=(\mathcal{A}-\lambda_j\mathcal{I})^{l_j}\alpha_j=0,$$

从而 $(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{l_j}=0$,于是$(\lambda-\lambda_j)^{l_j}$是 $\mathcal{A}|W_j$ 的一个零化多项式,因此 $\mathcal{A}|W_j$ 的最小多项式 $m_j(\lambda)=(\lambda-\lambda_j)^{t_j}$,其中 $t_j\leq l_j$.于是 $\mathcal{A}$ 的最小多项式 $m(\lambda)$ 为

$$\begin{aligned}m(\lambda)&=[(\lambda-\lambda_1)^{t_1},(\lambda-\lambda_2)^{t_2},\cdots,(\lambda-\lambda_s)^{t_s}]\\&=(\lambda-\lambda_1)^{t_1}(\lambda-\lambda_2)^{t_2}\cdots(\lambda-\lambda_s)^{t_s}.\end{aligned} \tag{8}$$

据唯一因式分解定理,从(8)式和(5)式得

$$t_1=l_1,\ t_2=l_2,\cdots,t_s=l_s.$$

因此 $\mathcal{A}|W_j$ 的最小多项式 $m_j(\lambda)=(\lambda-\lambda_j)^{l_j}$.

令 $\mathcal{B}_j=\mathcal{A}|W_j-\lambda_j\mathcal{I}$,$j=1,2,\cdots,s$.由于 $\mathcal{A}|W_j$ 的最小多项式 $m_j(\lambda)=(\lambda-\lambda_j)^{l_j}$,因此

$$(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{l_j}=0,\quad (\mathcal{A}|W_j-\lambda_j\mathcal{I})^{k_j}\neq 0,\text{当 } 0\leq k_j<l_j.$$

由此得出,$\mathcal{B}_j$ 是 W_j 上的幂零变换,且幂零指数为 l_j.由于 $\mathcal{A}|W_j$ 在 W_j 的上述基下的矩阵是 A_j,因此 $\mathcal{B}_j=\mathcal{A}|W_j-\lambda_j\mathcal{I}$ 在 W_j 的上述基下的矩阵 $B_j=A_j-\lambda_j I$.于是为了使 A_j 最简单,就应当使 B_j 最简单.由此提出一个问题:

设 $\mathcal{B}$ 是域 F 上 r 维线性空间 W 上的幂零变换,且幂零指数为 l,能否在 W 中找到一个基,使得 $\mathcal{B}$ 在此基下的矩阵最简单?

这个问题将在下一节详细讨论.

习　题　9.7

1. 求下列复矩阵的最小多项式:

(1) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$;　　(2) $\begin{pmatrix} 3 & 1 & -1 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$.

2. 求下列数域 K 上 Jordan 形矩阵的最小多项式:

(1) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$;　　(2) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$;

(3) $\begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$;　　(4) $\begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}$.

3. 第 2 题的第(3)、(4)小题的矩阵是否相似?

4. 第 1 题中的各个复矩阵是否可对角化?

5. 第 2 题中数域 K 上的各个 Jordan 形矩阵是否可对角化?

6. 证明:复数域上的周期矩阵一定可对角化.

7. 设数域 K 上的 n 级矩阵 A 满足

$$A^3 = 3A^2 + A - 3I,$$

判断 A 是否可对角化.

8. 设数域 K 上的 n 级矩阵 A 满足

$$A^3 = A^2 + 4A - 4I,$$

判断 A 是否可对角化.

9. 设 A 是有理数域 Q 上的 n 级非零矩阵,且 A 有一个零化多项式 $g(\lambda)$ 是 Q 上 r 次不可约多项式,$r>1$,判断 A 是否可对角化.如果把 A 看成复矩阵,那么 A 是否可对角化?

10. 设 V 是数域 K 上 3 维线性空间,V 上的线性变换 $\mathcal{A}$ 在基 $\alpha_1, \alpha_2, \alpha_3$ 下的矩阵 A 为

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & 0 & 1 \end{pmatrix}.$$

(1) 求 $\mathcal{A}$ 的最小多项式 $m(\lambda)$;

(2) 对应于 $m(\lambda)$的因式分解,写出 V 的直和分解式,并且求出分解式中出现的每个子空间的一个基.

11. 设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的线性变换,$\mathcal{A}$ 在 V 的一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵 A 为

$$A=\begin{pmatrix}0&0&\cdots&0&-a_0\\1&0&\cdots&0&-a_1\\0&1&\cdots&0&-a_2\\\cdots&\cdots&&\cdots&\cdots\\0&0&\cdots&1&-a_{n-1}\end{pmatrix},$$

求 A 的最小多项式 $m(\lambda)$和特征多项式 $f(x)$.

*12. 设 A 是域 F 上线性空间 V 上的线性变换,证明:如果 A 的最小多项式 $m(\lambda)$不可约,那么 $F[A]$是一个域.

*13. 证明:如果域 F 上的 n 级矩阵 A 与 B 都是可对角化的,并且 $AB=BA$,则存在域 F 上一个 n 级可逆矩阵 S,使得 $S^{-1}AS$ 与 $S^{-1}BS$ 都为对角矩阵.

§8 幂零变换的结构

在上一节的最后,我们指出,研究不可对角化的线性变换 A 的结构问题,如果 A 的最小多项式 $m(\lambda)$在 $F[\lambda]$中可分解成一次因式的乘积,那么可归结为研究幂零变换的结构.即研究域 F 上 r 维线性空间 W 上的幂零变换 B,能否在 W 中找到一个基,使得 B 在此基下的矩阵最简单?

命题1 设 B 是域 F 上 r 维线性空间 W 上的幂零变换,且幂零指数为 l,则存在 $\xi\in W$,使得向量组

$$B^{l-1}\xi,\ B^{l-2}\xi,\cdots,\ B\xi,\ \xi$$

线性无关,从而幂零指数 l 小于或等于 $\dim W$.

证明 由于 B 的幂零指数为 l,因此 $B^{l-1}\neq 0$.从而存在 $\xi\in W$,使得 $B^{l-1}\xi\neq 0$.于是 $B^{l-2}\xi,\cdots,B\xi,\xi$ 都不等于 0,而 $B^l\xi=0\xi=0$.据习题 9.1 第 9 题的结论得,$B^{l-1}\xi,B^{l-2}\xi,\cdots,B\xi,\xi$ 线性无关.从而 $l\le\dim W$. x

首先讨论幂零指数 $l=\dim W$ 的情形.此时据命题 1 得,存在 $\xi\in W$,使得 $B^{l-1}\xi,B^{l-2}\xi,\cdots,B\xi,\xi$ 线性无关.由于 $l=\dim W$,因此它们就是 W 的一个基.由于

$$B(B^{l-1}\xi)=B^l\xi=0,\ B(B^{l-2}\xi)=B^{l-1}\xi,\cdots,\ B(\xi)=B\xi,$$

因此 B 在 W 的基 $B^{l-1}\xi,B^{l-2}\xi,\cdots,B\xi,\xi$ 下的矩阵为

$$\begin{pmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\cdots&\cdots&\cdots&&\cdots\\0&0&0&\cdots&1\\0&0&0&\cdots&0\end{pmatrix}.\tag{1}$$

矩阵(1)是主对角元为0的 r 级Jordan块 $J_r(0)$.

其次讨论幂零指数 $l<\dim W$ 的情形.

此时,仍据命题1得,存在 $\xi\in W$,使得 $B^{l-1}\xi,\cdots,B\xi,\xi$ 线性无关,但是由于 $l<\dim W$,因此它们不是 W 的一个基.注意到它们生成的子空间 $\langle\xi,B\xi,\cdots,B^{l-1}\xi\rangle$ 是 B 的不变子空间,并且 B 在这个子空间上的限制在基 $B^{l-1}\xi,\cdots,B\xi,\xi$ 下的矩阵是一个Jordan块 $J_l(0)$(即,上述矩阵(1)).为此我们抽象出一个概念:

定义1　设 $\eta\in W$,如果存在正整数 t,使得

$$B^{t-1}\eta\neq 0,\text{而 } B^t\eta=0.$$

则子空间 $\langle\eta,B\eta,\cdots,B^{t-1}\eta\rangle$ 称为由 η 生成的 B-循环子空间,记作 $Z_t(\eta;B)$.

显然,$Z_t(\eta;B)$是 B 不变子空间.据习题9.1的第9题的结论得,$B^{t-1}\eta,\cdots,B\eta,\eta$ 线性无关,从而它们是 $Z_t(\eta;B)$的一个基.于是 t 是 $Z_t(\eta;B)$的维数.由前面的讨论知道,$B|Z_t(\eta;B)$在这个基下的矩阵是一个 t 级Jordan块 $J_t(0)$.于是只要我们能把 W 分解成若干个 B-循环子空间的直和,那么 B 的矩阵就是由若干个Jordan块组成的分块对角矩阵,即Jordan形矩阵.为了探讨这个问题,我们从另一个角度分析 $Z_t(\eta;B)$:

$$\begin{aligned}
&\alpha\in Z_t(\eta;B)\\
&\alpha=k_{t-1}B^{t-1}\eta+\cdots+k_1B\eta+k_0\eta\\
&\quad=(k_{t-1}B^{t-1}+\cdots+k_1B+k_0I)\eta\\
&\alpha=h(B)\eta,\quad h(B)=k_{t-1}B^{t-1}+\cdots+k_1B+k_0I.
\end{aligned}$$

因此

$$Z_t(\eta;B)=\{h(B)\eta\mid h(\lambda)\in F[\lambda],\deg h(\lambda)<t\}.\tag{2}$$

这表明 $Z_t(\eta;B)$是由 η 这样生成的:任取 $h(\lambda)\in F[\lambda]$,且 $\deg h(\lambda)<t$,η 在 $h(B)$下的象 $h(B)\eta$ 组成 $Z_t(\eta;B)$.由此受到启发,我们引出下述概念:

定义2　设 W 是域 F 上 r 维线性空间,B 是 W 上的一个线性变换,$\eta_1,\eta_2,\cdots,\eta_s$ 是 W 的一个向量组.对于 $\alpha\in W$,如果存在 $h_1(\lambda),h_2(\lambda),\cdots,h_s(\lambda)\in F[\lambda]$,使得

$$\alpha=h_1(B)\eta_1+h_2(B)\eta_2+\cdots+h_s(B)\eta_s,\tag{3}$$

则称 α 可以由 $\eta_1,\eta_2,\cdots,\eta_s$ $F[\lambda]$-表出,把 $h_i(\lambda)$称为 η_i 的多项式系数.如果 V 中每个向量 α 都可以由 $\eta_1,\eta_2,\cdots,\eta_s$ $F[\lambda]$-表出,则称 $\eta_1,\eta_2,\cdots,\eta_s$ 是 W 的一个 B-生成元系.

例如,W 的任意一个基 $\alpha_1,\alpha_2,\cdots,\alpha_r$ 就是 W 的一个 B-生成元系,这是因为对于任意 $\alpha\in W$,有

$$\begin{aligned}
\alpha&=k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r\\
&=(k_1I)\alpha_1+(k_2I)\alpha_2+\cdots+(k_rI)\alpha_r,
\end{aligned}$$

其中 $k_i \in F$,从而 $k_i \in F[\lambda]$, $i=1,2,\cdots,r$.

定义 3 设 B 是域 F 上 r 维线性空间 W 的线性变换.如果 W 有一个 B-生成元系 $\eta_1,\eta_2,\cdots,\eta_m$ 含 m 个向量,而 W 的任何 $m-1$ 个向量都不是 W 的 B-生成元系,则称 $\eta_1,\eta_2,\cdots,\eta_m$ 是 W 的一个最小 B-生成元系.

由于自然数集的任一非空子集必有最小数,因此 W 一定存在最小 B-生成元系. W 的最小 B-生成元系不唯一.

例如,设 B 是域 F 上 r 维线性空间 W 上的幂零变换,且幂零指数为 l,则存在 $\xi \in W$,使得 $B^{l-1}\xi \neq 0$.于是 ξ 生成的 B-循环子空间 $Z_l(\xi;B)$ 就有一个最小 $B|Z_l(\xi;B)$-生成元系:ξ.

利用最小 B-生成元系,可以得出下面结果:

定理 2 设 B 是域 F 上有限维线性空间 W 上的幂零变换,则 W 能分解成一些 B-循环子空间的直和.

证明 对于 W 的最小 B-生成元系所含向量的个数 s 作数学归纳法.

$s=1$ 时,设 η 是 W 的一个最小 B-生成元系.由于 W 是有限维线性空间,因此存在一个正整数 t,使得 $B^{t-1}\eta \neq 0$,而 $B^t\eta=0$.由于 η 是 W 的 B-生成元系,因此

$$
\begin{aligned}
W &= \{h(B)\eta \mid h(\lambda) \in F[\lambda], \deg h(\lambda) < t\} \\
&= Z_t(\eta;B).
\end{aligned}
$$

假设 $s=m-1$ 时,命题为真,来看 $s=m$ 的情形.设 W 的一个最小 B-生成元系为 $\eta_1,\eta_2,\cdots,\eta_m$.形如

$$
h_1(B)\eta_1 + h_2(B)\eta_2 + \cdots + h_m(B)\eta_m = 0 \tag{4}
$$

的一个等式称为生成元的一个关系.如果 $h_i(\lambda)=0$, $i=1,2,\cdots,m$,则称这个关系是平凡的;否则称为非平凡的. W 的 B-生成元系必有非平凡的关系(例如,设 $B^{t-1}\eta_1 \neq 0$, $B^t\eta_1=0$,则 $B^t\eta_1+0\eta_2+\cdots+0\eta_m=0$,其中 $h_1(\lambda)=\lambda^t \neq 0$,因此这是一个非平凡关系).在 W 的所有最小 B-生成元系的全部关系中,生成元的非零多项式系数组成的集合必有首项系数为 1 的次数最低的多项式(因为自然数集的非空子集必有最小数),设为 $f_1(\lambda)$.不妨设 $f_1(\lambda)$ 是 W 的最小 B-生成元系 $\eta_1,\eta_2,\cdots,\eta_m$ 中 η_1 的多项式系数,即 $\eta_1,\eta_2,\cdots,\eta_m$ 有一个关系:

$$
f_1(B)\eta_1 + h_2(B)\eta_2 + \cdots + h_m(B)\eta_m = 0. \tag{5}
$$

我们断言 $f_1(\lambda) \mid h_2(\lambda)$.理由如下:设

$$
h_2(\lambda) = g(\lambda)f_1(\lambda) + r(\lambda), \quad \deg r(\lambda) < \deg f_1(\lambda). \tag{6}
$$

则由(5)式得

$$
\begin{aligned}
0 &= f_1(B)\eta_1 + [g(B)f_1(B) + r(B)]\eta_2 + \cdots + h_m(B)\eta_m \\
&= f_1(B)[\eta_1 + g(B)\eta_2] + r(B)\eta_2 + \cdots + h_m(B)\eta_m.
\end{aligned} \tag{7}
$$

由于 W 中任一向量 α 可表示成

$$\begin{aligned}\alpha &= p_1(B)\eta_1 + p_2(B)\eta_2 + \cdots + p_m(B)\eta_m \\ &= p_1(B)[\eta_1 + g(B)\eta_2] + [p_2(B) - p_1(B)g(B)]\eta_2 + \cdots + p_m(B)\eta_m,\end{aligned}$$

因此 $\eta_1 + g(B)\eta_2, \eta_2, \cdots, \eta_m$ 也是 W 的一个最小 B-生成元系.于是据 $f_1(\lambda)$ 的选择,从(7)式得出,$r(\lambda) = 0$.从而 $f_1(\lambda) | h_2(\lambda)$.

同理可证 $f_1(\lambda) | h_j(\lambda), j = 3, \cdots, m$.设

$$h_j(\lambda) = g_j(\lambda) f_1(\lambda), \quad j = 3, \cdots, m,$$

则(5)式可以写成

$$f_1(B)\eta_1 + g(B)f_1(B)\eta_2 + g_3(B)f_1(B)\eta_3 + \cdots + g_m(B)f_1(B)\eta_m = 0. \quad (8)$$

令

$$\xi_1 = \eta_1 + g(B)\eta_2 + g_3(B)\eta_3 + \cdots + g_m(B)\eta_m. \quad (9)$$

从(9)式得,η_1 可以由 $\xi_1, \eta_2, \cdots, \eta_m$ $F[\lambda]$-表出,因此 $\xi_1, \eta_2, \cdots, \eta_m$ 也是 W 的一个最小 B-生成元系.从(8)式得,

$$f_1(B)\xi_1 = 0. \quad (10)$$

从 $f_1(\lambda)$ 的选择知道,$f_1(\lambda)$ 是使 $f_1(B)\xi_1 = 0$ 的次数最低的首项系数为1的多项式.又由于对于 ξ_1 存在正整数 k_1,使得 $B^{k_1-1}\xi_1 \neq 0$, $B^{k_1}\xi_1 = 0$.因此 $f_1(\lambda) = \lambda^{k_1}$. ξ_1 生成一个 k_1 维 B-循环子空间 $Z_{k_1}(\xi_1; B)$.令

$$U = \{h_2(B)\eta_2 + \cdots + h_m(B)\eta_m \mid h_j(\lambda) \in F[\lambda], j = 2, \cdots, m\}, \quad (11)$$

容易看出,U 是 W 的子空间,且 U 是 B 不变子空间.从而 $B|U$ 是 U 的幂零变换,由于 $\xi_1, \eta_2, \cdots, \eta_m$ 是 W 的一个最小 B-生成元系,因此

$$W = Z_{k_1}(\xi_1; B) + U. \quad (12)$$

任取 $\gamma \in Z_{k_1}(\xi_1; B) \cap U$,则存在 $q_i(\lambda) \in F[\lambda], i = 1, 2, \cdots, m$,且 $\deg q_1(\lambda) < k_1$,使得

$$\gamma = q_1(B)\xi_1 = q_2(B)\eta_2 + \cdots + q_m(B)\eta_m, \quad (13)$$

即

$$q_1(B)\xi_1 - q_2(B)\eta_2 - \cdots - q_m(B)\eta_m = 0. \quad (14)$$

由于 $\deg q_1(\lambda) < k_1 = \deg f_1(\lambda)$,因此假如 $q_1(\lambda) \neq 0$,则(14)式与 $f_1(\lambda)$ 的选择矛盾.从而 $q_1(\lambda) = 0$.由此得出,$\gamma = q_1(B)\xi_1 = 0\xi_1 = 0$.于是 $Z_{k_1}(\xi_1, B) \cap U = 0$.所以

$$W = Z_{k_1}(\xi_1; B) \oplus U. \quad (15)$$

$\eta_2, \cdots, \eta_m$ 必为 U 的一个最小 B-生成元系(否则,U 有小于 $m-1$ 个向量组成的最小 B-生成元系,从而由(12)式知道,W 有小于 m 个向量组成的最小

$\mathcal{B}$ - 生成元系,矛盾). 据归纳假设, U 可以分解成一些 $\mathcal{B}|U$ - 循环子空间的直和:

$$U = Z_{k_2}(\xi_2;\mathcal{B}|U) \oplus \cdots \oplus Z_{k_m}(\xi_m;\mathcal{B}|U). \tag{16}$$

显然有 $Z_{k_j}(\xi_j;\mathcal{B}|U) = Z_{k_j}(\xi_j;\mathcal{B})$, $j = 2,\cdots,m$. 因此

$$W = Z_{k_1}(\xi_1;\mathcal{B}) \oplus Z_{k_2}(\xi_2;\mathcal{B}) \oplus \cdots \oplus Z_{k_m}(\xi_m;\mathcal{B}). \tag{17}$$

根据数学归纳法原理,命题得证. $\square$

有了定理 2,我们就可以了解幂零变换的结构了. 即我们有下述定理:

定理 3　设 $\mathcal{B}$ 是域 F 上 r 维线性空间 W 上的幂零变换,其幂零指数为 l. 则 W 中存在一个基,使得 $\mathcal{B}$ 在此基下的矩阵 B 为 Jordan 形矩阵,其中每个 Jordan 块的主对角元都是 0,且级数不超过 l, Jordan 块的总数等于 $\mathcal{B}$ 的特征子空间 W_0 的维数, t 级 Jordan 块的个数 $N(t)$ 为

$$N(t) = \operatorname{rank} B^{t+1} + \operatorname{rank} B^{t-1} - 2\operatorname{rank} B^{t}; \tag{18}$$

这个 Jordan 形矩阵 B 称为幂零变换 $\mathcal{B}$ 的 Jordan 标准形,除去 Jordan 块的排列次序外, $\mathcal{B}$ 的 Jordan 标准形是唯一的.

证明　由定理 2 得到

$$W = \bigoplus_{i=1}^{m} Z_{k_i}(\xi_i;\mathcal{B}). \tag{19}$$

由于 $\mathcal{B}^l = 0$, 因此 $k_i \le l$, $i = 1,2,\cdots,m$. 在 $Z_{k_i}(\xi_i;\mathcal{B})$ 中取一个基 $\mathcal{B}^{k_i-1}\xi_i,\cdots,\mathcal{B}\xi_i,\xi_i$, $i = 1,2,\cdots,m$, 它们合起来成为 W 的一个基. $\mathcal{B}$ 在此基下的矩阵 B 为

$$B = \operatorname{diag}\{J_{k_1}(0), J_{k_2}(0), \cdots, J_{k_m}(0)\}, \tag{20}$$

这是 Jordan 形矩阵.

主对角元为 0 的每个 Jordan 块有且只有一个零行,因此 B 中 Jordan 块的总数 m 等于 B 中零行的数目,后者等于 r 减去 B 中非零行的数目. 而 B 中非零行的数目等于 $\operatorname{rank} B$, 因此 $m = r - \operatorname{rank} B$. 此式右端等于齐次线性方程组 $BX = 0$ 的解空间的维数,由于幂零变换 $\mathcal{B}$ 的特征值为 0,因此方程组 $(0I - B)X = 0$ 的解空间的维数等于 $\mathcal{B}$ 的特征子空间 W_0 的维数. 从而

$$m = \dim W_0. \tag{21}$$

对于 n 级 Jordan 块 $J_n(0)$, $\operatorname{rank} J_n(0) = n - 1$. 直接计算得,当 $1 < s < n$ 时,

$$J_n(0)^s=\begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ \cdots & & \cdots & \cdots & \cdots & \cdots & & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \cdots & & \cdots & \cdots & \cdots & \cdots & & \cdots & \cdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \quad (\text{上方 } s\text{个}, \text{右侧 } s\text{个}) \tag{22}$$

因此当 $1<s<n$ 时, $\mathrm{rank}\,J_n(0)^s=n-s$. 显然, 当 $s\geq n$ 时, $J_n(0)^s=0$. 由此看出, Jordan 块 $J_n(0)$ 的方幂的秩与级数 n 有关系. 因此为了求 t 级 Jordan 块的个数 $N(t)$, 我们先来求 $\mathrm{rank}\,B^t$, $1\leq t\leq l$,

$$B^t=\mathrm{diag}\{J_{k_1}(0)^t, J_{k_2}(0)^t, \cdots, J_{k_m}(0)^t\}. \tag{23}$$

当 $k_i\leq t$ 时, $J_{k_i}(0)^t=0$. 因此为了计算 $\mathrm{rank}\,B^t$, 只要考虑级数大于 t 的 Jordan 块的 t 次幂, 这种 Jordan 块的 t 次幂的秩等于级数减 t, 于是

$$\begin{aligned}\mathrm{rank}\,B^t&=[(t+1)-t]N(t+1)+[(t+2)-t]N(t+2)+\cdots+(l-t)N(l)\\ &=N(t+1)+2N(t+2)+\cdots+(l-t)N(l).\end{aligned} \tag{24}$$

当 $2\leq t\leq l$ 时, 从(24)式得

$$\mathrm{rank}\,B^{t-1}=N(t)+2N(t+1)+\cdots+(l-t+1)N(l) \tag{25}$$

当 $t=1$ 时, (25)式左端 $=\mathrm{rank}\,B^0=\mathrm{rank}\,I=r$, (25)式右端为

$$N(1)+2N(2)+\cdots+lN(l)=r,$$

因此(25)式仍然成立.

把(25)式减去(24)式得

$$\mathrm{rank}\,B^{t-1}-\mathrm{rank}\,B^t=N(t)+N(t+1)+\cdots+N(l). \tag{26}$$

当 $1\leq t<l$ 时, 把(26)式的 t 换成 $t+1$ 得

$$\mathrm{rank}\,B^t-\mathrm{rank}\,B^{t+1}=N(t+1)+N(t+2)+\cdots+N(l). \tag{27}$$

于是当 $1\leq t<l$ 时, 把(26)式减去(27)式得

$$\mathrm{rank}\,B^{t-1}+\mathrm{rank}\,B^{t+1}-2\mathrm{rank}\,B^t=N(t) \tag{28}$$

当 $t=l$ 时, (28)式左端 $=\mathrm{rank}\,B^{l-1}=[l-(l-1)]N(l)=N(l)$, 右端 $=N(l)$, 因此(28)式对于 $1\leq t\leq l$ 都成立. 从(28)式得

$$N(t)=\mathrm{rank}\,B^{t+1}+\mathrm{rank}\,B^{t-1}-2\mathrm{rank}\,B^t.$$

由于 B 中 Jordan 块的主对角元都为 0, Jordan 块的总数等于 B 的特征子空间的维数, t 级 Jordan 块的个数 $N(t)$ 由 B^{t+1}, B^{t-1}, B^t 的秩决定. 因此除去 Jordan 块的排列次序外, B 的 Jordan 标准形是唯一的. x

从定理 3 立即得到下述结论:

推论 4 设 B 是域 F 上的幂零矩阵,其幂零指数为 l.则 B 相似于一个 Jordan 形矩阵,其中每个 Jordan 块的主对角元为 0,且级数不超过 l,Jordan 块的总数等于 B 的特征子空间的维数(即齐次线性方程组 $BX=0$ 的解空间的维数),t 级 Jordan 块的个数 $N(t)$ 为

$$N(t)=\text{rank}\,B^{t+1}+\text{rank}\,B^{t-1}-2\text{rank}\,B^{t};\qquad(29)$$

这个 Jordan 形矩阵称为 B 的 Jordan 标准形,除去 Jordan 块的排列次序外,B 的 Jordan 标准形是唯一的. x

习 题 9.8

1. 设数域 K 上的 4 级矩阵 B 为

$$\begin{pmatrix} -1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix},$$

(1) 说明 B 是幂零矩阵,求 B 的幂零指数;

(2) 求 B 的 Jordan 标准形.

2. 证明:如果域 F 上的 n 级矩阵 B 是幂零矩阵,则对一切正整数 k,有 $\text{tr}(B^k)=0$.

3. 证明:n 级复矩阵 B 是幂零矩阵当且仅当 B 的特征值全是 0.

§9 线性变换的 Jordan 标准形

现在我们利用幂零变换的结构来研究最小多项式可以分解成一次因式乘积的线性变换的结构.

定理 1 设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的线性变换,如果 $\mathcal{A}$ 的最小多项式 $m(\lambda)$在 $F[\lambda]$中能分解成一次因式的乘积:

$$m(\lambda)=(\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\cdots(\lambda-\lambda_s)^{l_s},\qquad(1)$$

则 V 中存在一个基,使得 $\mathcal{A}$ 在此基下的矩阵 A 为 Jordan 形矩阵,其主对角元为 $\mathcal{A}$ 的全部特征值;主对角元为 λ_j 的 Jordan 块的总数 N_j 为

$$N_j=\dim V-\text{rank}(A-\lambda_j I),\quad j=1,2,\cdots,s,\qquad(2)$$

且其中每个 Jordan 块的级数不超过 l_j,t 级 Jordan 块 $J_t(\lambda_j)$的个数 $N_j(t)$为

$$N_j(t)=\text{rank}(A-\lambda_j I)^{t+1}+\text{rank}(A-\lambda_j I)^{t-1}-2\text{rank}(A-\lambda_j I)^{t};\qquad(3)$$

这个 Jordan 形矩阵 A 称为 $\mathcal{A}$ 的 Jordan 标准形,除去 Jordan 块的排列次序外,$\mathcal{A}$

的 Jordan 标准形是唯一的.

证明　由(1)式看出, $\lambda_1,\lambda_2,\cdots,\lambda_s$ 是 $\mathcal{A}$ 的所有不同的特征值.由(1)式得出

$$V=\operatorname{Ker} m(\mathcal{A})=\operatorname{Ker}(\mathcal{A}-\lambda_1\mathcal{I})^{l_1}\quad\cdots\quad\operatorname{Ker}(\mathcal{A}-\lambda_s\mathcal{I})^{l_s}. \tag{4}$$

令 $W_j=\operatorname{Ker}(\mathcal{A}-\lambda_j\mathcal{I})^{l_j}$, $j=1,2,\cdots,s$.则 W_j 是 $\mathcal{A}$ 的不变子空间, $j=1,2,\cdots,s$. 由(4)式得

$$V=W_1\quad W_2\quad\cdots\quad W_s. \tag{5}$$

令 $\mathcal{B}_j=\mathcal{A}|W_j-\lambda_j\mathcal{I}$,在§7 的最后已证 $\mathcal{B}_j$ 是 W_j 上的幂零变换,其幂零指数为 l_j.于是据§8 的定理 3 得,在 W_j 中存在一个基,使得 $\mathcal{B}_j$ 在此基下的矩阵 $\boldsymbol{B}_j$ 是 Jordan 形矩阵.从而 $\mathcal{A}|W_j$ 在 W_j 的这个基下的矩阵 $\boldsymbol{A}_j=\boldsymbol{B}_j+\lambda_j\boldsymbol{I}$ 是主对角元为 λ_j 的 Jordan 形矩阵.把 $W_j(j=1,2,\cdots,s)$的上述基合起来便成为 V 的一个基, $\mathcal{A}$ 在此基下的矩阵 $\boldsymbol{A}$ 为

$$\boldsymbol{A}=\operatorname{diag}\{\boldsymbol{A}_1,\boldsymbol{A}_2,\cdots,\boldsymbol{A}_s\}. \tag{6}$$

于是 $\boldsymbol{A}$ 是 Jordan 形矩阵,其主对角元是 $\mathcal{A}$ 的全部特征值.

由于 $\mathcal{B}_j$ 的幂零指数为 l_j,因此 $\boldsymbol{B}_j$ 中每个 Jordan 块的级数不超过 l_j(据§8 的定理 3),从而 $\boldsymbol{A}_j$ 中每个 Jordan 块的级数不超过 l_j.当 $1\leq t\leq l_j$ 时, $\boldsymbol{A}_j$ 中 t 级 Jordan 块的个数 $N_j(t)$等于 $\boldsymbol{B}_j$ 中 t 级 Jordan 块的个数,因此

$$\begin{aligned}
N_j(t)&=\operatorname{rank}\boldsymbol{B}_j^{t+1}+\operatorname{rank}\boldsymbol{B}_j^{t-1}-2\operatorname{rank}\boldsymbol{B}_j^{t}\\
&=\operatorname{rank}(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{t+1}+\operatorname{rank}(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{t-1}\\
&\quad-2\operatorname{rank}(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{t}\\
&=\ \dim W_j-\dim\operatorname{Ker}(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{t+1}\\
&\quad+\ \dim W_j-\dim\operatorname{Ker}(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{t-1}\\
&\quad-2\ \dim W_j-\dim\operatorname{Ker}(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{t}\\
&=2\dim\operatorname{Ker}(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{t}-\dim\operatorname{Ker}(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{t+1}\\
&\quad-\dim\operatorname{Ker}(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{t-1}.
\end{aligned}\tag{7}$$

当 $i\leq l_j$ 时,有

$$\begin{aligned}
&\alpha\in\operatorname{Ker}(\mathcal{A}-\lambda_j\mathcal{I})^{i}\\
&(\mathcal{A}-\lambda_j\mathcal{I})^{i}\alpha=0\\
&\alpha\in\operatorname{Ker}(\mathcal{A}-\lambda_j\mathcal{I})^{l_j}=W_j\text{且}(\mathcal{A}-\lambda_j\mathcal{I})^{i}\alpha=0\\
&\alpha\in W_j\text{且}(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{i}\alpha=0\\
&\alpha\in\operatorname{Ker}(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{i},
\end{aligned}\tag{8}$$

因此
$$\operatorname{Ker}(\mathcal{A}-\lambda_j\mathcal{I})^{i}=\operatorname{Ker}(\mathcal{A}|W_j-\lambda_j\mathcal{I})^{i}. \tag{9}$$

当 $i=l_j+1$ 时,(9)式仍然成立(理由见后面的注).于是从(7)式和(9)式得

$$N_j(t)=2\dim\operatorname{Ker}(\mathcal{A}-\lambda_j\mathcal{I})^{t}-\dim\operatorname{Ker}(\mathcal{A}-\lambda_j\mathcal{I})^{t+1}$$

$$
\begin{aligned}
&- \dim \mathrm{Ker}(A-\lambda_j I)^{t-1} \\
&= \mathrm{rank}(A-\lambda_j I)^{t+1} + \mathrm{rank}(A-\lambda_j I)^{t-1} \\
&- 2\mathrm{rank}(A-\lambda_j I)^{t}. \qquad (10)
\end{aligned}
$$

A_j 中 Jordan 块的总数 N_j 等于 B_j 中 Jordan 块的总数,从而等于 B_j 的特征子空间$(W_j)_0$ 的维数.由于

$$\alpha\in(W_j)_0 \qquad B_j\alpha=0\alpha=0 \qquad \alpha\in \mathrm{Ker}\, B_j,$$

因此$(W_j)_0=\mathrm{Ker}\, B_j$.从而

$$
\begin{aligned}
N_j &= \dim(W_j)_0 = \dim \mathrm{Ker}\, B_j = \dim \mathrm{Ker}(A|W_j - \lambda_j I) \\
&= \dim \mathrm{Ker}(A-\lambda_j I) = \dim V - \mathrm{rank}(A-\lambda_j I). \qquad (11)
\end{aligned}
$$

由于 A 的 Jordan 标准形 A 中,主对角线上元素是 A 的全部特征值,主对角元为特征值 λ_j 的 Jordan 块总数 N_j 由 V 的维数和 $\mathrm{rank}(A-\lambda_j I)$ 决定,主对角元为 λ_j 的 t 级 Jordan 块的个数 $N_j(t)$ 由$(A-\lambda_j I)^{t+1}$,$(A-\lambda_j I)^{t-1}$,$(A-\lambda_j I)^{t}$的秩决定,因此除去 Jordan 块的排列次序外,A 的 Jordan 标准形是唯一的. x

注:当 $i=l_j+1$ 时,(9)式仍然成立的理由如下:从(1)式得

$$m(\lambda)(\lambda-\lambda_j)=(\lambda-\lambda_1)^{l_1}\cdots(\lambda-\lambda_j)^{l_j+1}\cdots(\lambda-\lambda_s)^{l_s}. \qquad (12)$$

于是 V 可以分解成

$$V=\mathrm{Ker}(A-\lambda_1 I)^{l_1} \quad \cdots \quad \mathrm{Ker}(A-\lambda_j I)^{l_j+1} \quad \cdots \quad \mathrm{Ker}(A-\lambda_s I)^{l_s}. \qquad (13)$$

任取 $\beta\in\mathrm{Ker}(A-\lambda_j I)^{l_j+1}$,据(4)式,$\beta$ 可以分解成

$$\beta=\beta_1+\cdots+\beta_j+\cdots+\beta_s, \qquad (14)$$

其中 $\beta_u\in\mathrm{Ker}(A-\lambda_u I)^{l_u}$, $u=1,2,\cdots,s$.由于 $\beta_j\in\mathrm{Ker}(A-\lambda_j I)^{l_j}$,因此 $\beta_j\in\mathrm{Ker}(A-\lambda_j I)^{l_j+1}$.从而 $\beta_j-\beta\in\mathrm{Ker}(A-\lambda_j I)^{l_j+1}$.从(14)式得

$$0=\beta_1+\cdots+(\beta_j-\beta)+\cdots+\beta_s. \qquad (15)$$

(15)式是 0 在 V 的直和分解式(13)中的一种表示方式.由于表法唯一,因此 $\beta_j-\beta=0$.从而 $\beta=\beta_j\in\mathrm{Ker}(A-\lambda_j I)^{l_j}$.由此得出,$\mathrm{Ker}(A-\lambda_j I)^{l_j+1} \quad \mathrm{Ker}(A-\lambda_j I)^{l_j}$.反过来的包含关系是显然的.因此

$$\mathrm{Ker}(A-\lambda_j I)^{l_j+1}=\mathrm{Ker}(A-\lambda_j I)^{l_j}. \qquad (16)$$

现在任取 $\eta\in\mathrm{Ker}(A|W_j-\lambda_j I)^{l_j+1}$,则 $\eta\in W_j$.

由于 $W_j=\mathrm{Ker}(A-\lambda_j I)^{l_j}=\mathrm{Ker}(A|W_j-\lambda_j I)^{l_j}$,因此

$\eta\in\mathrm{Ker}(A|W_j-\lambda_j I)^{l_j}$.由此得出

$$\mathrm{Ker}(A|W_j-\lambda_j I)^{l_j+1}=\mathrm{Ker}(A|W_j-\lambda_j I)^{l_j}. \qquad (17)$$

从(16)、(17)和(9)式得

$$\mathrm{Ker}(A-\lambda_j I)^{l_j+1}=\mathrm{Ker}(A|W_j-\lambda_j I)^{l_j+1}. \qquad (18)$$

从定理1立即得到下述结论:

推论2　设 A 是域 F 上的 n 级矩阵,如果 A 的最小多项式 $m(\lambda)$ 在 $F[\lambda]$ 中能分解成一次因式的乘积:

$$m(\lambda)=(\lambda-\lambda_1)^{l_1}(\lambda-\lambda_2)^{l_2}\cdots(\lambda-\lambda_s)^{l_s}, \tag{19}$$

则 A 相似于一个Jordan形矩阵,其主对角线上元素是 A 的全部特征值;主对角元为 λ_j 的Jordan块总数 N_j 为

$$N_j=n-\text{rank}(A-\lambda_j I),\ j=1,2,\cdots,s, \tag{20}$$

且其中每个Jordan块的级数不超过 l_j, t 级Jordan块的个数 $N_j(t)$ 为

$$N_j(t)=\text{rank}(A-\lambda_j I)^{t+1}+\text{rank}(A-\lambda_j I)^{t-1}-2\text{rank}(A-\lambda_j I)^{t}; \tag{21}$$

这个Jordan形矩阵称为 A 的Jordan标准形,除去Jordan块的排列次序外, A 的Jordan标准形是唯一的.　x

由于复数域上每个次数大于0的一元多项式都可以分解成一次因式的乘积,因此复数域上有限维线性空间的每一个线性变换都有Jordan标准形,从而复数域上每一个方阵都有Jordan标准形.

例1　求复数域上下述矩阵的约当标准形:

$$A=\begin{pmatrix}2&3&2\\1&8&2\\-2&-14&-3\end{pmatrix}.$$

解　A 的特征多项式 $f(\lambda)$ 是

$$|\lambda I-A|=\begin{vmatrix}\lambda-2&-3&-2\\-1&\lambda-8&-2\\2&14&\lambda+3\end{vmatrix}=(\lambda-1)(\lambda-3)^2,$$

于是 A 的全部特征值是1,3(二重).

对于特征值 $\lambda_1=1$,它是 $f(\lambda)$ 的1重根,因此它在 A 的约当标准形的主对角线上只出现1次.

对于特征值 $\lambda_2=3$,先求 $\text{rank}(A-3I)$:

$$A-3I=\begin{pmatrix}-1&3&2\\1&5&2\\-2&-14&-6\end{pmatrix}\rightarrow\begin{pmatrix}-1&3&2\\0&8&4\\0&0&0\end{pmatrix}.$$

因此 $\text{rank}(A-3I)=2$.从而主对角元为3的约当块的总数为

$$N(\lambda_2)=3-2=1.$$

综上述得, A 的约当标准形为 $\begin{pmatrix}1&0&0\\0&3&1\\0&0&3\end{pmatrix}$.

由于域 F 上的 n 级矩阵 A 如果有 Jordan 标准形,则除了 Jordan 块的排列次序外, A 的 Jordan 标准形是唯一的.因此我们可以引进下述概念:

定义 1 域 F 上的 n 级矩阵 A 如果有 Jordan 标准形 J,则把 J 中所有 Jordan块的最小多项式组成的有重集叫做 A 的初等因子组,简称为 A 的初等因子.

例如,设

$$A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix},$$

则 A 的初等因子是: $(\lambda-3)^2, \lambda-5, \lambda-5$.

命题 3 域 F 上两个 Jordan 块相等当且仅当它们的最小多项式相等.

证明 必要性是显然的.

充分性.设两个 Jordan 块的最小多项式都是 $(\lambda- a)^t$,则它们的主对角元都是 a,级数都是 t.因此它们都是

$$\begin{pmatrix} a & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & a & 1 & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & a & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & a \end{pmatrix}_{t\times t}.$$

x

定理 4 两个 n 级复矩阵 A 与 B 相似当且仅当它们的初等因子相同.

证明 设复矩阵 A 的 Jordan 标准形为 J.由于相似关系具有对称性和传递性,因此 A 与 B 相似当且仅当 J 也是 B 的 Jordan 标准形,从而 A 与 B 的 Jordan 标准形中所有 Jordan 块对应相等.再据命题 3 得,后者当且仅当 A 与 B 的 Jordan标准形中所有 Jordan 块的最小多项式对应相等,即 A 与 B 的初等因子相同. x

从定理 4 看出,在复数域上的所有 n 级矩阵组成的集合 $M_n(C)$ 中,初等因子是相似关系下的一组完全不变量.

我们还可以把 λ - 矩阵 $\lambda I - A$ 通过初等变换化成对角矩阵,来求复矩阵 A 的初等因子.[①]

定义 2 V 上线性变换 A 的一个 Jordan 基是 V 的一个基,它使得 A 在这个基下的矩阵为 Jordan 形矩阵.

---

① 参看丘维声编著《高等代数(下册)》(高等教育出版社 1996 年 12 月第 1 版)的阅读材料七的例 1.

当我们已经求出 $\mathcal{A}$ 的Jordan标准形 J 以后,为了求出 $\mathcal{A}$ 的一个Jordan基,只要把原来的基到Jordan基的过渡矩阵 S 求出即可.由于 $J=S^{-1}AS$,所以 S 是矩阵方程

$$AX=XJ \tag{22}$$

的解并且应为可逆矩阵.如果 $\dim V=n$,则(22)是 n^2 个未知量 $x_{ij}(i,j=1,\cdots,n)$的由 n^2 个方程组成的线性方程组,解这个线性方程组,可求出 $X=(x_{ij})$,选取可逆矩阵(因为 A 的 Jordan 标准形存在,所以满足方程(22)的可逆矩阵一定存在),便可作为过渡矩阵 S.

例2　求例1的线性变换 $\mathcal{A}$ 的一个Jordan基.

解　设 $X=(X_1,X_2,X_3)$,由 A 的Jordan标准形以及方程(22)得

$$A(X_1,X_2,X_3)=(X_1,3X_2,X_2+3X_3), \tag{23}$$

所以
$$AX_1=X_1,\quad AX_2=3X_2,\quad AX_3=X_2+3X_3.$$

由此看出,X_1 是 A 的属于1的一个特征向量,解方程组

$$(I-A)Y=0,$$

求得 $X_1=(2,0,-1)'$.同理,X_2 是 A 的属于3的一个特征向量,解方程组

$$(3I-A)Y=0,$$

求得 $X_2=(1,-1,2)'$.再去解方程组$(A-3I)Y=X_2$,

$$\begin{pmatrix} -1 & 3 & 2 & 1 \\ 1 & 5 & 2 & -1 \\ -2 & -14 & -6 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -\frac{1}{2} & -1 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

它的一个特解是 $Y_0=(-1,0,0)'$.取 $X_3=(-1,0,0)'$,则

$$X=\begin{pmatrix} 2 & 1 & -1 \\ 0 & -1 & 0 \\ -1 & 2 & 0 \end{pmatrix}. \tag{24}$$

容易看出 X 是可逆矩阵.它就可作为 V 的原来的基 $\alpha_1,\alpha_2,\alpha_3$ 到 $\mathcal{A}$ 的一个Jordan基 ξ_1,ξ_2,ξ_3 的过渡矩阵.所以

$$(\xi_1,\xi_2,\xi_3)=(\alpha_1,\alpha_2,\alpha_3)\begin{pmatrix} 2 & 1 & -1 \\ 0 & -1 & 0 \\ -1 & 2 & 0 \end{pmatrix},$$

即 $\mathcal{A}$ 的一个Jordan基是:

$$\xi_1=2\alpha_1-\alpha_3,\quad \xi_2=\alpha_1-\alpha_2+2\alpha_3,\quad \xi_3=-\alpha_1.$$

矩阵的Jordan标准形的最重要的应用之一是计算矩阵的多项式,譬如,要

计算 A^m,其中 m 是很大的正整数.如果 A 有 Jordan 标准形 J,并且用上述方法求出了可逆矩阵 X,使得 $X^{-1}AX = J$,那么

$$A^m = (XJX^{-1})^m = XJ^m X^{-1}. \tag{25}$$

每个 Jordan 块 $J_t(\lambda_j)$ 的方幂易于计算:

$$J_t(\lambda_j)^r = (\lambda_j I + J_t(0))^r. \tag{26}$$

由于 $\lambda_j I$ 与 $J_t(0)$ 可交换,因此可用二项式定理展开(26)式右端.注意:当 $r \geq t$ 时,$J_t(0)^r = 0$.这表明,计算 J^m 比直接计算 A^m 要简便得多.而且一旦把 J 和 X 求出来后,对一切正整数 m,可以用同一个公式(25)计算 A^m.

例 3 对于例 1 中的 A,计算 A^{10}.

解 从例 1 知道,A 的 Jordan 标准形是

$$J = \operatorname{diag}\left\{(1), \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}\right\}.$$

从例 2 知道,(24)式给出的矩阵 X 使得 $X^{-1}AX = J$.从而 $A^m = XJ^m X^{-1}$.易求出

$$X^{-1} = \begin{pmatrix} 0 & -2 & -1 \\ 0 & -1 & 0 \\ -1 & -5 & -2 \end{pmatrix}.$$

当 $r \geq 2$,有

$$\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}^r = \left(3I + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right)^r$$

$$= 3^r I + r3^{r-1} I \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 3^r & r3^{r-1} \\ 0 & 3^r \end{pmatrix}.$$

所以

$$A^{10} = X \operatorname{diag}\left\{(1)^{10}, \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}^{10}\right\} X^{-1}$$

$$= \begin{pmatrix} 2 & 1 & -1 \\ 0 & -1 & 0 \\ -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3^{10} & 10\cdot 3^9 \\ 0 & 0 & 3^{10} \end{pmatrix} \begin{pmatrix} 0 & -2 & -1 \\ 0 & -1 & 0 \\ -1 & -5 & -2 \end{pmatrix}$$

$$= \begin{pmatrix} -7\cdot 3^9 & -38\cdot 3^9 - 4 & -14\cdot 3^9 - 2 \\ 10\cdot 3^9 & 53\cdot 3^9 & 20\cdot 3^9 \\ -20\cdot 3^9 & -106\cdot 3^9 + 2 & -40\cdot 3^9 + 1 \end{pmatrix}.$$

习 题 9.9

1. 求下列数域 K 上的矩阵的 Jordan 标准形:

(1) $\begin{pmatrix} 0 & 1 & 0 \\ -4 & 4 & 0 \\ -2 & 1 & 2 \end{pmatrix}$;　(2) $\begin{pmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{pmatrix}$;

(3) $\begin{pmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{pmatrix}$;　(4) $\begin{pmatrix} 1 & -3 & 4 \\ 4 & -7 & 8 \\ 6 & -7 & 7 \end{pmatrix}$;

(5) $\begin{pmatrix} 3 & -4 & 0 & 2 \\ 4 & -5 & -2 & 4 \\ 0 & 0 & 3 & -2 \\ 0 & 0 & 2 & -1 \end{pmatrix}$;　(6) $\begin{pmatrix} 3 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 3 & 0 & 5 & -3 \\ 4 & -1 & 3 & -1 \end{pmatrix}$.

2. 对于第 1 题各个小题中的矩阵,写出它的最小多项式.

3. 证明: $J_t(a) \sim J_t(a)'$.

4. 证明:任一 n 级复矩阵 A 与它的转置 A′相似.

5. 证明:设 n 级复矩阵 A 的 n 个特征值是 $\lambda_1, \lambda_2, \cdots, \lambda_n$,则对于任一正整数 m, A^m 的 n 个特征值是 $\lambda_1^m, \lambda_2^m, \cdots, \lambda_n^m$.

6. 证明:设 n 级复矩阵 A 的 n 个特征值是 $\lambda_1, \lambda_2, \cdots, \lambda_n$,则对于任一复系数多项式 $g(x)$,复矩阵 $g(A)$的 n 个特征值是 $g(\lambda_1), g(\lambda_2), \cdots, g(\lambda_n)$.

7. 设

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

(1) 证明:当 $k \geq 3$ 时,有

$$A^k = A^{k-2} + A^2 - \mathbf{I};$$

(2) 求 A^{100}.

§10　线性函数与对偶空间

在本章§1 中我们曾指出,函数的定积分把闭区间[a, b]上的每一个连续函数 $f(x)$对应到一个实数: $\int_a^b f(x)dx$;并且这个映射保持加法与数量乘法.

矩阵的迹把域 F 上每一个 n 级矩阵 $A = (a_{ij})$对应到 F 中一个元素 $\sum_{i=1}^n a_{ii}$,并且有

$$\mathbf{tr}(A + B) = \mathbf{tr}(A) + \mathbf{tr}(B), \quad \forall A, B \in M_n(F),$$
$$\mathbf{tr}(kA) = k\mathbf{tr}(A), \quad \forall A \in M_n(F), k \in F.$$

定义 1　设 V 是域 F 上的一个线性空间, V 到 F 的一个映射 f 如果满足

$$f(\alpha+\beta)=f(\alpha)+f(\beta),\quad \forall\ \alpha,\beta\in V,$$
$$f(k\alpha)=kf(\alpha),\quad \forall\ \alpha\in V,\ k\in F,$$

则称 f 是 V 上的一个线性函数.

由于域 F 可以看成是自身上的线性空间,因此 V 上的线性函数也就是域 F 上线性空间 V 到线性空间 F 的线性映射.从而本章§1至§3关于线性映射的结果对于线性函数也成立.

线性函数是十分重要的一类函数.下面再举几个例子.

例1　给定域 F 中 n 个元素 $a_1,a_2,\cdots,a_n$,令

$$f:\quad F^n\qquad F$$
$$(x_1,x_2,\cdots,x_n)\longmapsto a_1x_1+a_2x_2+\cdots+a_nx_n,$$

容易直接验证 f 是 F^n 上的一个线性函数.

例2　令

$$f:\quad Z_2^2\qquad Z_2$$
$$(x_1,x_2)\longmapsto x_1^2+x_2^2,$$

判别 f 是不是 Z_2^2 上的一个线性函数.

解　任取 $(x_1,x_2),(y_1,y_2)\in Z_2^2$,有

$$\begin{aligned}f\ (x_1,x_2)+(y_1,y_2)\ &=f(x_1+y_1,x_2+y_2)\\ =(x_1+y_1)^2+(x_2+y_2)^2&=x_1^2+y_1^2+x_2^2+y_2^2\\ =f(x_1,x_2)+f(y_1,y_2).&\end{aligned}$$

$$f\ 1\cdot(x_1,x_2)\ =f(x_1,x_2)=1\cdot f(x_1,x_2),$$
$$f\ 0\cdot(x_1,x_2)\ =f(0,0)=0=0\cdot f(x_1,x_2).$$

因此 f 是 Z_2^2 上的一个线性函数.

设 V 是域 F 上 n 维线性空间,f 是 V 上的一个线性函数.由于 f 是线性空间 V 到 F 的线性映射,因此 f 完全被它在 V 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 上的作用所决定.即只要知道 $f(\alpha_1),f(\alpha_2),\cdots,f(\alpha_n)$,就可以求出 V 中任一向量 $\alpha=\sum_{i=1}^{n}x_i\alpha_i$ 在 f 下的象:

$$f(\alpha)=\sum_{i=1}^{n}x_if(\alpha_i).\tag{1}$$

(1)式称为线性函数 f 在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的表达式.

反之,在 V 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$,在 F 中任意取定 n 个元素 $a_1,a_2,\cdots,a_n$.对于 V 中任一向量 $\alpha=\sum_{i=1}^{n}x_i\alpha_i$,规定

$$f(\alpha)=\sum_{i=1}^{n}x_i a_i,\tag{2}$$

据§1的定理1得，f 是 V 上的一个线性函数，它满足：

$$f(\alpha_i)=a_i,\quad i=1,2,\cdots,n.\tag{3}$$

并且满足(3)式的 V 上的线性函数是唯一的.

设 V 是域 F 上的一个线性空间，由于 V 上的线性函数就是线性空间 V 到 F 的线性映射，因此据§1有关线性映射的结论得，V 上所有线性函数组成的集合 $\mathrm{Hom}(V,F)$ 是域 F 上的一个线性空间，称它是 V 上的线性函数空间.如果 V 的维数为 n，则

$$\dim \mathrm{Hom}(V,F)=\dim V\cdot\dim F=n\cdot 1=n\tag{4}$$

从而

$$\mathrm{Hom}(V,F)\cong V.\tag{5}$$

在 V 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，我们按照(2)式和(3)式指出的方法构造 V 上的 n 个线性函数：给定 $i\in\{1,2,\cdots,n\}$，令

$$f_i(\alpha_j)=\begin{cases}1, & 当\ j=i,\\ 0, & 当\ j\neq i,\end{cases}\tag{6}$$

$$f_i\left(\sum_{j=1}^{n}x_j\alpha_j\right)=\sum_{j=1}^{n}x_j f_i(\alpha_j)=x_i,\tag{7}$$

则 f_i 是 V 上的一个线性函数.于是我们得到 V 上的 n 个线性函数 $f_1,f_2,\cdots,f_n$.设

$$k_1f_1+k_2f_2+\cdots+k_nf_n=0,$$

则

$$(k_1f_1+k_2f_2+\cdots+k_nf_n)\alpha_j=0,\quad j=1,2,\cdots,n$$

根据线性.映射的加法和纯量乘法运算的定义得

$$k_1f_1(\alpha_j)+\cdots+k_jf_j(\alpha_j)+\cdots+k_nf_n(\alpha_j)=0,\quad j=1,2,\cdots,n.$$

从而得出

$$k_j=0,\quad j=1,2,\cdots,n.$$

因此 $f_1,f_2,\cdots,f_n$ 线性无关.由于 $\dim \mathrm{Hom}(V,F)=n$，因此 $f_1,f_2,\cdots,f_n$ 是 $\mathrm{Hom}(V,F)$ 的一个基.我们把由(6)式和(7)式定义的 $f_1,f_2,\cdots,f_n$ 称为基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的对偶基，把 V 上的线性函数空间 $\mathrm{Hom}(V,F)$ 称为 V 的对偶空间，简记作 V^*.

设 V 是域 F 上的一个 n 维线性空间，V 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$，它的对偶基 $f_1,f_2,\cdots,f_n$ 是 V^* 的一个基.任取 $\alpha=\sum_{j=1}^{n}x_j\alpha_j\in V$，从(7)式得

$$f_i(\alpha)=f_i\left(\sum_{j=1}^{n}x_j\alpha_j\right)=x_i.\tag{8}$$

(8)式表明,α在基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的坐标的第 i 个分量等于对偶基的第 i 个线性函数 f_i 在 α 处的函数值 $f_i(\alpha)$.任取 $f=\sum_{i=1}^{n} k_i f_i \in V^*$,我们有

$$f(\alpha_j)=\sum_{i=1}^{n} k_i f_i(\alpha_j)=k_j. \tag{9}$$

(9)式表明,线性函数 f 在α_j 处的函数值 $f(\alpha_j)$等于 f 在对偶基 $f_1,f_2,\cdots,f_n$ 下的坐标的第 j 个分量.

定理 1 设 V 是域 F 上的一个 n 维线性空间,在 V 中取两个基:$\alpha_1,\alpha_2,\cdots,\alpha_n$ 与 $\beta_1,\beta_2,\cdots,\beta_n$,它们的对偶基分别是:$f_1,f_2,\cdots,f_n$ 与 $g_1,g_2,\cdots,g_n$.如果 V 中基$\alpha_1,\alpha_2,\cdots,\alpha_n$ 到基$\beta_1,\beta_2,\cdots,\beta_n$ 的过渡矩阵是 A,则 V^* 中基 $f_1,f_2,\cdots,f_n$ 到基 $g_1,g_2,\cdots,g_n$ 的过渡矩阵 B 为

$$B=(A^{-1})'. \tag{10}$$

证明 由$(\beta_1,\beta_2,\cdots,\beta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)A$ 得

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\beta_1,\beta_2,\cdots,\beta_n)A^{-1}. \tag{11}$$

从(11)式看出,α_j 在基$\beta_1,\beta_2,\cdots,\beta_n$ 下的坐标是 A^{-1}的第 j 列,从而坐标的第 i 个分量为 $A^{-1}(i;j)$.由于 $\beta_1,\beta_2,\cdots,\beta_n$ 的对偶基是 $g_1,g_2,\cdots,g_n$,因此根据(8)式表明的结论得,α_j 在基$\beta_1,\beta_2,\cdots,\beta_n$ 下的坐标的第 i 个分量等于 $g_i(\alpha_j)$.从而 $A^{-1}(i;j)=g_i(\alpha_j)$.

由于 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的对偶基是 $f_1,f_2,\cdots,f_n$,因此据(9)式表明的结论得,$g_i(\alpha_j)$等于 g_i 在 $f_1,f_2,\cdots,f_n$ 下的坐标的第 j 个分量.由已知条件知道

$$(g_1,g_2,\cdots,g_n)=(f_1,f_2,\cdots,f_n)B, \tag{12}$$

从(12)式看出,g_i 在 $f_1,f_2,\cdots,f_n$ 下的坐标的第 j 个分量等于 $B(j;i)$.因此 $g_i(\alpha_j)=B(j;i)$.综上所述得

$$A^{-1}(i;j)=g_i(\alpha_j)=B(j;i)=B'(i;j),$$

其中 $i,j=1,2,\cdots,n$.因此 $A^{-1}=B'$.从而

$$B=(A^{-1})'.$$ X

设 V 是域 F 上的一个 n 维线性空间,则 $V \cong V^*$.在 V 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$,它的对偶基 $f_1,f_2,\cdots,f_n$ 是 V^* 的一个基.令

$$\sigma:\quad V \qquad\qquad V^*$$

$$\alpha=\sum_{i=1}^{n} x_i\alpha_i \longmapsto \sum_{i=1}^{n} x_i f_i, \tag{13}$$

据第 8 章§3 的定理 1 的证明过程知道,σ是 V 到 V^* 的一个同构映射.记 $\sigma(\alpha)=f_\alpha$.对于 $\beta=\sum_{i=1}^{n} y_i\alpha_i \in V$,据(8)式有

$$f_\alpha(\beta) = \left(\sum_{i=1}^{n} x_i f_i\right)\beta = \sum_{i=1}^{n} x_i f_i(\beta) = \sum_{i=1}^{n} x_i y_i. \tag{14}$$

(14)式表明,α 在 σ 下的象 f_α 在 β 处的函数值 $f_\alpha(\beta)$ 等于 α 与 β 的坐标的对应分量乘积之和.

上述讨论是对于域 F 上任一 n 维线性空间进行的,因此对于域 F 上 n 维线性空间 V 的对偶空间 V^*,也可以考虑它的对偶空间 $(V^*)^*$,简记成 V^{**}. 称 V^{**} 是 V 的双重对偶空间. 我们有

$$V \cong V^* \cong V^{**}. \tag{15}$$

设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 的一个基,它的对偶基是 $f_1, f_2, \cdots, f_n$. 据(13)式, V 到 V^* 有一个同构映射 σ_1, V 中向量 $\alpha = \sum_{i=1}^{n} x_i\alpha_i$ 在 σ_1 下的象记作 f_α. 同理,有 V^* 到 V^{**} 的一个同构映射 σ_2, f_α 在 σ_2 下的象记作 α^{**}. 任取 $f \in V^*$, 据(14)式表明的结论得, $\alpha^{**}(f)$ 等于 f_α 与 f 在基 $f_1, f_2, \cdots, f_n$ 下的坐标的对应分量乘积之和. 据(13)以及(9)式表明的结论得

$$f_\alpha = \sum_{i=1}^{n} x_i f_i, \quad f = \sum_{i=1}^{n} f(\alpha_i) f_i, \tag{16}$$

因此

$$\alpha^{**}(f) = \sum_{i=1}^{n} x_i f(\alpha_i) = f\left(\sum_{i=1}^{n} x_i\alpha_i\right) = f(\alpha). \tag{17}$$

令 $\tau = \sigma_2\sigma_1$, 则 τ 是 V 到 V^{**} 的一个同构映射,并且

$$\tau(\alpha) = \sigma_2(\sigma_1(\alpha)) = \sigma_2(f_\alpha) = \alpha^{**}, \quad \forall\, \alpha \in V, \tag{18}$$

从(17)式知道

$$\alpha^{**}(f) = f(\alpha), \quad \forall\, f \in V^*. \tag{19}$$

从(18)式和(19)式看到, α 在 V 到 V^{**} 的同构映射 τ 下的象 α^{**} 不依赖于 V 中基的选择. 我们称这样的同构映射为标准同构或自然同构.

由于 V 到 V^{**} 存在自然同构 τ,因此可以把 α 与 $\tau(\alpha) = \alpha^{**}$ 等同,从而可以把 V 与 V^{**} 等同. 于是可以把 V 看成 V^* 的对偶空间. 这样 V 与 V^* 就互为对偶空间. 这就是为什么把 V^* 称为 V 的对偶空间的原因.

习　题　9.10

1. 设 $V = C[a, b]$, V 上的下列函数哪些是线性函数?

(1) $f(x) \longmapsto \int_a^b f^2(x)\,dx$;

(2) $f(x) \longmapsto \int_a^b f(x)g(x)dx$,

其中 $g(x)$是一个固定的函数.

2. 设 V 是域 F 上的一个 3 维线性空间,$\alpha_1,\alpha_2,\alpha_3$ 是 V 的一个基,f 是 V 上的一个线性函数,已知

$$f(\alpha_1+2\alpha_3)=4,\ f(\alpha_2+3\alpha_3)=0,\ f(4\alpha_1+\alpha_2)=5,$$

求 $f(x_1\alpha_1+x_2\alpha_2+x_3\alpha_3)$.

3. V 及$\alpha_1,\alpha_2,\alpha_3$ 同第 2 题,试找出一个线性函数 f,使

$$f(3\alpha_1+\alpha_2)=2,\ f(\alpha_2-\alpha_3)=1,\ f(2\alpha_1+\alpha_3)=2.$$

*4. 设 V 是域 F 上一个线性空间,char $F=0$.设 $f_1,f_2,\cdots,f_s$ 都是 V 上的线性函数,并且它们都不是零函数.证明:存在 $\alpha\in V$,使得

$$f_i(\alpha)\neq 0,\ i=1,2,\cdots,s.$$

5. 设 V 是数域 K 上一个 3 维线性空间,$\alpha_1,\alpha_2,\alpha_3$ 是 V 的一个基,它的对偶基是 f_1,f_2,f_3.设

$$\beta_1=2\alpha_1+\alpha_2+2\alpha_3,\ \beta_2=\alpha_1+2\alpha_2-2\alpha_3,\ \beta_3=-2\alpha_1+2\alpha_2+\alpha_3,$$

证明:β_1,β_2,β_3 是 V 的一个基,并且求它的对偶基(用 f_1,f_2,f_3 表出).

6. 设 $V=\mathbf{R}[x]_3$,对于 $g(x)\in V$,定义

$$f_1(g(x))=\int_0^1 g(x)dx,\ f_2(g(x))=\int_0^2 g(x)dx,\ f_3(g(x))=\int_0^{-1} g(x)dx.$$

证明:f_1,f_2,f_3 是 V^* 的一个基;并且求出 V 的一个基$g_1(x),g_2(x),g_3(x)$,使得 f_1,f_2,f_3 是它的对偶基.

7. 设 $\mathcal{A}$ 是域 F 上 n 维线性空间 V 上的一个线性变换.

(1) 证明:对于 $f\in V^*$,有 $f\mathcal{A}\in V^*$;

(2) 定义 V^* 到自身的一个映射 $\mathcal{A}^*$ 为

$$f\longmapsto f\mathcal{A},$$

证明 $\mathcal{A}^*$ 是 V^* 上的线性变换;

(3) 设 $\alpha_1,\cdots,\alpha_n$ 是 V 的一个基,$f_1,\cdots,f_n$ 是它的对偶基,并且 $\mathcal{A}$ 在$\alpha_1,\cdots,\alpha_n$ 下的矩阵为 A,证明:$\mathcal{A}^*$ 在 $f_1,\cdots,f_n$ 下的矩阵为 A'(因此 $\mathcal{A}^*$ 称为 $\mathcal{A}$ 的转置映射).

8. 设 V 是域 F 上一个线性空间,

(1) $f_1,\cdots,f_s$ 是 V 上 s 个线性函数,证明下述集合

$$W=\{\alpha\in V\mid f_i(\alpha)=0,i=1,\cdots,s\}$$

是 V 的一个子空间.W 称为线性函数 $f_1,\cdots,f_s$ 的零化子空间.

*(2) 证明:若 V 是有限维的,则 V 的任一子空间都是某些线性函数的零化子空间.

第 10 章　具有度量的线性空间

迄今为止,我们对于线性空间和线性映射的研究都是围绕线性空间的加法与纯量乘法两种运算来进行的.在上册第 4 章 §6 中,我们对于实数域上的 n 维向量空间 R^n 定义了一个内积,从而在 R^n 中引进了长度、正交等度量概念.由此受到启发,本章我们来讨论如何分别在实数域、复数域、任意域上的线性空间中引进度量概念,然后研究具有度量的线性空间的结构,并且研究保持度量的线性变换的性质.

§1　双线性函数

我们已经知道,R^n 上的一个内积(α,β)是 R^n 上的一个二元实值函数,并且它具有下列性质:

1° $(\alpha,\beta)=(\beta,\alpha)$(对称性);

2° $(\alpha+\gamma,\beta)=(\alpha,\beta)+(\gamma,\beta)$(线性);

3° $(k\alpha,\beta)=k(\alpha,\beta)$(线性);

4° $(\alpha,\alpha)\geq 0$,等号成立当且仅当 $\alpha=0$(正定性).

由 1°,2°,3°,可以得出

$$(k_1\alpha_1+k_2\alpha_2,\beta)=k_1(\alpha_1,\beta)+k_2(\alpha_2,\beta),$$

$$(\alpha,k_1\beta_1+k_2\beta_2)=k_1(\alpha,\beta_1)+k_2(\alpha,\beta_2).$$

从 R^n 上的内积满足的上述两个式子,我们抽象出域 F 上线性空间 V 上的双线性函数的概念:

定义 1　设 V 是域 F 上的一个线性空间,$V\times V$ 到 F 的一个映射 f 如果满足:$\forall\ \alpha_1,\alpha_2,\beta_1,\beta_2,\alpha,\beta\in V$,$k_1,k_2\in F$,有

(i) $f(k_1\alpha_1+k_2\alpha_2,\beta)=k_1f(\alpha_1,\beta)+k_2f(\alpha_2,\beta)$;

(ii) $f(\alpha,k_1\beta_1+k_2\beta_2)=k_1f(\alpha,\beta_1)+k_2f(\alpha,\beta_2)$,

则称 f 是 V 上的一个双线性函数,f 也可写成 $f(\alpha,\beta)$.

条件(i)表明:当 β 固定时,映射 $\alpha\longmapsto f(\alpha,\beta)$ 是 V 上的一个线性函数,记作 β_R;

条件(ii)表明:当 α 固定时,映射 $\beta\longmapsto f(\alpha,\beta)$ 是 V 上的一个线性函数,

记作 α_L.

例 1 设 $V = M_n(F)$,令

$$f(A, B) = \mathrm{tr}(AB), \quad \forall A, B \in V,$$

容易验证 f 是 V 上的一个双线性函数.

例 2 设 $V = C[a, b]$,令

$$f(g(x), h(x)) = \int_a^b g(x)h(x)dx, \quad \forall g(x), h(x) \in V,$$

容易验证 f 是 V 上的一个双线性函数.

例 3 设 F 是任一域,在 F^n 中,对于

$$\alpha = (a_1, a_2, \cdots, a_n), \quad \beta = (b_1, b_2, \cdots, b_n),$$

令

$$f(\alpha, \beta) = \sum_{i=1}^{n} a_i b_i,$$

容易验证 f 是 F^n 上的一个双线性函数.

设 V 是域 F 上 n 维线性空间, f 是 V 上的一个双线性函数. V 中取一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 设 α, β 在此基下的坐标分别为 $X = (x_1, x_2, \cdots, x_n)'$, $Y = (y_1, y_2, \cdots, y_n)'$,则

$$f(\alpha, \beta) = f\left(\sum_{i=1}^{n} x_i \alpha_i, \sum_{j=1}^{n} y_j \alpha_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j f(\alpha_i, \alpha_j). \tag{1}$$

令

$$A = \begin{pmatrix} f(\alpha_1, \alpha_1) & f(\alpha_1, \alpha_2) & \cdots & f(\alpha_1, \alpha_n) \\ f(\alpha_2, \alpha_1) & f(\alpha_2, \alpha_2) & \cdots & f(\alpha_2, \alpha_n) \\ \cdots & \cdots & & \cdots \\ f(\alpha_n, \alpha_1) & f(\alpha_n, \alpha_2) & \cdots & f(\alpha_n, \alpha_n) \end{pmatrix}, \tag{2}$$

称 A 是双线性函数 f 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的度量矩阵,它是由 f 及基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 唯一决定的.

从(1)式得

$$f(\alpha, \beta) = X'AY. \tag{3}$$

(3)式和(1)式都是双线性函数 f 在基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的表达式.

容易看出,如果对任意 $X, Y \in F^n$,都有 $X'AY = X'BY$,则 $A = B$(利用 $\varepsilon_i' A \varepsilon_j = a_{ij}$,即可证得).

定理 1 设 f 是域 F 上 n 维线性空间 V 上的一个双线性函数, V 中取两个基: $\alpha_1, \alpha_2, \cdots, \alpha_n$ 与 $\beta_1, \beta_2, \cdots, \beta_n$. 设

$$(\beta_1, \beta_2, \cdots, \beta_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)P, \tag{4}$$

f 在这两个基下的度量矩阵分别为 A, B,则

$$B = P'AP. \tag{5}$$

证明　设 $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n) X = (\beta_1, \beta_2, \cdots, \beta_n) X_0$, $\beta = (\alpha_1, \alpha_2, \cdots, \alpha_n) Y = (\beta_1, \beta_2, \cdots, \beta_n) Y_0$,则

$$X = PX_0, \quad Y = PY_0,$$

从而 $$f(\alpha, \beta) = X'AY = (PX_0)'A(PY_0) = X_0'(P'AP)Y_0.$$

又有 $f(\alpha, \beta) = X_0'BY_0$,由此得出, $X_0'BY_0 = X_0'(P'AP)Y_0$,其中 X_0, Y_0 是 K^n 中任意向量,因此 $B = P'AP$. x

定理 1 表明, V 上的双线性函数 f 在不同基下的度量矩阵是合同的. 由于合同的矩阵有相同的秩,因此我们把双线性函数 f 在 V 的一个基下的度量矩阵的秩称为 f 的矩阵秩,记作 $\text{rank}_m f$.

定义 2　设 f 是域 F 上线性空间 V 上的一个双线性函数, V 的下述子集

$$\{\alpha \in V \mid f(\alpha, \beta) = 0, \forall \beta \in V\} \tag{6}$$

称为 f 在 V 中的左根,记作 $\text{rad}_L V$; V 的另一个子集

$$\{\beta \in V \mid f(\alpha, \beta) = 0, \forall \alpha \in V\} \tag{7}$$

称为 f 在 V 中的右根,记作 $\text{rad}_R V$.

容易看出, f 的左根、右根都是 V 的子空间.

例 4　在 R^3 中,设 $\alpha = (x_1, x_2, x_3)'$, $\beta = (y_1, y_2, y_3)'$. 令

$$f(\alpha, \beta) = x_1 y_1 - x_3 y_3 = (x_1, x_2, x_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

容易验证, f 是 R^3 上的一个双线性函数. 令 $\alpha_1 = (0, 1, 0)$,则对于 R^3 中任意 $\beta = (y_1, y_2, y_3)'$,都有

$$f(\alpha_1, \beta) = 0 \cdot y_1 - 0 \cdot y_3 = 0.$$

因此 α_1 属于 f 的左根.

定义 3　如果 V 上的双线性函数 f 的左根和右根都是零子空间,则称 f 是非退化的.

定理 2　设 f 是域 F 上 n 维线性空间 V 上的一个双线性函数, f 在基 α_1, $\alpha_2, \cdots, \alpha_n$ 下的度量矩阵为 A,则 f 是非退化的当且仅当 A 是满秩矩阵.

证明　先证 f 的左根为 0 当且仅当 A 满秩. 设

$$\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n) X, \quad \beta = (\alpha_1, \alpha_2, \cdots, \alpha_n) Y,$$

则 $f(\alpha, \beta) = X'AY$,于是

$$f \text{ 的左根 } \text{rad}_L(V) = 0$$

从 $f(\alpha,\beta)=0$, $\forall\ \beta\in V$ 可以推出 $\alpha=0$

从 $X'AY=0$, $\forall\ Y\in F^n$ 可以推出 $X=0$

从 $X'A\varepsilon_i=0, i=1,2,\cdots,n$ 可以推出 $X=0$

从 $X'A(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=0$ 可以推出 $X=0$

从 $X'AI=0$ 可以推出 $X=0$

从 $A'X=0$ 可以推出 $X=0$

齐次线性方程组 $A'X=0$ 只有零解

$\mathrm{rank}(A')=n$

$\mathrm{rank}(A)=n$.

同理可证, f 的右根为 0 当且仅当 A 满秩. 因此 f 非退化当且仅当它的度量矩阵 A 满秩. x

从定理 2 的证明中还可得出, f 的左根等于 0 当且仅当 f 的右根等于 0.

定义 4 设 f 是域 F 上线性空间 V 上的一个双线性函数, 如果

$$f(\alpha,\beta)=f(\beta,\alpha),\quad \forall\ \alpha,\beta\in V, \tag{8}$$

则称 f 是对称的; 如果

$$f(\alpha,\beta)=-f(\beta,\alpha),\quad \forall\ \alpha,\beta\in V, \tag{9}$$

则称 f 是斜对称的(或反对称的).

实数域上线性空间 R^n 上的内积 (α,β) 是对称双线性函数. 例 1、例 2、例 3 中的双线性函数也都是对称的.

例 5 在 R^2 中, 对于

$$\alpha=(x_1,x_2),\beta=(y_1,y_2),$$

令

$$f(\alpha,\beta)=x_1y_2-x_2y_1.$$

容易验证 f 是斜对称双线性函数.

设双线性函数 f 在 V 的基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的度量矩阵为 A, 如果 f 是对称的, 则

$$f(\alpha_i,\alpha_j)=f(\alpha_j,\alpha_i),\quad i,j=1,2,\cdots,n,$$

因此 A 是对称矩阵. 反之, 如果 A 是对称矩阵, 则对任意 $\alpha=(\alpha_1,\cdots,\alpha_n)X$, $\beta=(\alpha_1,\cdots,\alpha_n)Y$, 有

$$\begin{aligned}f(\alpha,\beta)&=X'AY=(X'AY)'=Y'A'(X')'\\&=Y'AX=f(\beta,\alpha),\end{aligned}$$

因此 f 是对称的. 这样我们证明了:

双线性函数 f 是对称的 f 的度量矩阵是对称矩阵.

同理可证:

双线性函数 f 是斜对称的 $\Leftrightarrow$ f 的度量矩阵是斜对称矩阵，

定理 3　设 f 是数域 K 上 n 维线性空间 V 上的一个对称双线性函数,则 V 中存在一个基,使得 f 在此基下的度量矩阵是对角矩阵.

证明　任取 V 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$,设 f 在这个基下的度量矩阵为 A,则 A 是对称矩阵.由于任一数域 K 上的对称矩阵一定合同于一个对角矩阵,因此存在 K 上可逆矩阵 C,使得 $C'AC=D$,其中 D 为对角矩阵.令

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)C,$$

由于 C 是可逆的,因此 $\eta_1,\eta_2,\cdots,\eta_n$ 也是 V 的一个基.f 在基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的度量矩阵为

$$C'AC=D.$$ x

定理 3 的结论可以推广到特征不为 2 的任意域上的有限维线性空间中,即我们有下述定理 4.

定理 4　设 f 是特征不为 2 的域 F 上的 n 维线性空间 V 上的对称双线性函数,则 V 中存在一个基,使 f 在此基下的度量矩阵为对角矩阵.

*证明　对维数 n 作归纳法.

$n=1$ 时,f 在任一个基下的矩阵是 1 级矩阵,这是对角矩阵.

假设 F 上 $n-1$ 维线性空间上的对称双线性函数有此性质,现在看 n 维线性空间 V 上的对称双线性函数 f.

假如对一切 $\alpha,\beta\in V$,都有 $f(\alpha,\beta)=0$,则结论显然成立.下面设 $f(\alpha,\beta)$ 不全为零.这时一定存在一个向量 $\alpha_1\neq 0$,使得

$$f(\alpha_1,\alpha_1)\neq 0.$$

否则,若对一切 $\alpha\in V$,都有 $f(\alpha,\alpha)=0$,则对任意 $\alpha,\beta\in V$,有

$$\begin{aligned}0&=f(\alpha+\beta,\alpha+\beta)=f(\alpha,\alpha+\beta)+f(\beta,\alpha+\beta)\\&=f(\alpha,\alpha)+f(\alpha,\beta)+f(\beta,\alpha)+f(\beta,\beta)\\&=2f(\alpha,\beta).\end{aligned}$$

因为 $\mathrm{char}\,F\neq 2$,所以从上式得 $f(\alpha,\beta)=0$,矛盾.记 $f(\alpha_1,\alpha_1)=d_1$,把 α_1 扩充成 V 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$.用类似于 Schimidt 正交化方法,令

$$\tilde{\alpha}_i=\alpha_i-\frac{f(\alpha_1,\alpha_i)}{f(\alpha_1,\alpha_1)}\alpha_1,\quad i=2,\cdots,n, \tag{10}$$

直接计算得

$$f(\alpha_1,\tilde{\alpha}_i)=0,\quad i=2,\cdots,n. \tag{11}$$

从(10)式可得出 $\alpha_1,\tilde{\alpha}_2,\cdots,\tilde{\alpha}_n$ 仍是 V 的一个基.令

$$W=\langle\tilde{\alpha}_2,\cdots,\tilde{\alpha}_n\rangle,$$

则 $V=\langle\alpha_1\rangle\oplus W$,把 f 看成 W 上的双线性函数,它仍是对称的.据归纳假设,W

中存在一个基 $\eta_2,\cdots,\eta_n$,使得 f 在这个基下的度量矩阵为对角矩阵 D_1,即

$$f(\eta_i,\eta_j)=0,\ i\neq j,\ i,j=2,\cdots,n. \tag{12}$$

由(11)式得出,对任意 $\beta\in W$,有 $f(\alpha_1,\beta)=0$.从而 $f(\alpha_1,\eta_i)=0$, $i=2,\cdots,n$.于是 f 在 V 的一个基$\alpha_1,\eta_2,\cdots,\eta_n$ 下的度量矩阵为对角矩阵 $D=\mathrm{diag}\{d_1,D_1\}$.

据归纳法原理,对一切自然数 n,命题成立. x

从定理 4 得出,如果 V 上的对称双线性函数是非退化的,则存在 V 的一个基 $\eta_1,\cdots,\eta_n$,使得

$$f(\eta_i,\eta_i)\neq 0,\ i=1,\cdots,n,$$

$$f(\eta_i,\eta_j)=0,\ i\neq j,\ i,j=1,\cdots,n.$$

定理 4 的意义之一在于可简化计算对称双线性函数在任意一对向量上的函数值.由于一定存在 V 的一个基 $\eta_1,\eta_2,\cdots,\eta_n$,使得 f 在此基下的度量矩阵为对角矩阵 $\mathrm{diag}\{d_1,\cdots,d_n\}$,从而对任意

$$\alpha=x_1\eta_1+\cdots+x_n\eta_n,\ \beta=y_1\eta_1+\cdots+y_n\eta_n,$$

有

$$f(\alpha,\beta)=d_1x_1y_1+d_2x_2y_2+\cdots+d_nx_ny_n. \tag{13}$$

现在我们来讨论斜对称双线性函数.设 f 是域 F 上 n 维线性空间 V 的斜对称双线性函数.

*如果 $\mathrm{char}\ F=2$,则对任意 $\alpha,\beta\in V$,有 $2f(\alpha,\beta)=0$.从而 $f(\alpha,\beta)+f(\alpha,\beta)=0$.于是 $f(\alpha,\beta)=-f(\alpha,\beta)$.由于 f 是斜对称的,所以

$$f(\alpha,\beta)=-f(\beta,\alpha)=f(\beta,\alpha).$$

这表明 f 又是对称的.因此,在特征为 2 的域上的线性空间中,斜对称双线性函数与对称双线性函数是一致的.

下面设 $\mathrm{char}\ F\neq 2$.这时对任意 $\alpha\in V$,有 $f(\alpha,\alpha)=-f(\alpha,\alpha)$.从而 $2f(\alpha,\alpha)=0$.由此得出 $f(\alpha,\alpha)=0$.

定理 5 设 f 是特征不为 2 的域 F 上的 n 维线性空间 V 上的斜对称双线性函数,则存在 V 的一个基 $\varepsilon_1,\varepsilon_{-1},\cdots,\varepsilon_r,\varepsilon_{-r},\eta_1,\cdots,\eta_s$,使得 f 在这个基下的度量矩阵具有形式

$$\mathrm{diag}\left\{\begin{pmatrix}0&1\\-1&0\end{pmatrix},\cdots,\begin{pmatrix}0&1\\-1&0\end{pmatrix},0,\cdots,0\right\}, \tag{14}$$

即

$$\begin{aligned}&f(\varepsilon_i,\varepsilon_{-i})=1, && i=1,\cdots,r,\\&f(\varepsilon_i,\varepsilon_j)=0, && i+j\neq 0,\\&f(\alpha,\eta_k)=0, && \alpha\in V,\ k=1,\cdots,s.\end{aligned} \tag{15}$$

*证明 对维数 n 作归纳法.

$n=1$ 时, f 的度量矩阵为 0.

假设维数小于 n 时,命题成立.来看 n 维情形.

如果 $f=0$,则 f 的度量矩阵为 0.

下面设 $f\neq 0$.这时 V 中一定存在线性无关的向量 ε_1,α_2,使得 $f(\varepsilon_1,\alpha_2)\neq 0$.理由是:假如对一切线性无关的向量 α,β,都有

$$f(\alpha,\beta)=0.$$

又对一切线性相关的向量 $\alpha, k\alpha$,有

$$f(\alpha,k\alpha)=kf(\alpha,\alpha)=0.$$

则 $f=0$,矛盾.令 $\varepsilon_{-1}=f(\varepsilon_1,\alpha_2)^{-1}\alpha_2$,则 $f(\varepsilon_1,\varepsilon_{-1})=1$.显然 $\varepsilon_1,\varepsilon_{-1}$ 仍线性无关.把 $\varepsilon_1,\varepsilon_{-1}$ 扩充成 V 的一个基 $\varepsilon_1,\varepsilon_{-1},\beta_3,\cdots,\beta_n$.令

$$\beta_i'=\beta_i-f(\beta_i,\varepsilon_{-1})\varepsilon_1+f(\beta_i,\varepsilon_1)\varepsilon_{-1},\quad i=3,\cdots,n, \tag{16}$$

则

$$\begin{aligned}f(\varepsilon_1,\beta_i')&=f(\varepsilon_1,\beta_i)-f(\beta_i,\varepsilon_{-1})f(\varepsilon_1,\varepsilon_1)\\&\quad+f(\beta_i,\varepsilon_1)f(\varepsilon_1,\varepsilon_{-1})\\&=f(\varepsilon_1,\beta_i)+f(\beta_i,\varepsilon_1)=0,\ i=3,\cdots,n,\end{aligned} \tag{17}$$

$$\begin{aligned}f(\varepsilon_{-1},\beta_i')&=f(\varepsilon_{-1},\beta_i)-f(\beta_i,\varepsilon_{-1})f(\varepsilon_{-1},\varepsilon_1)\\&\quad+f(\beta_i,\varepsilon_1)f(\varepsilon_{-1},\varepsilon_{-1})\\&=f(\varepsilon_{-1},\beta_i)+f(\beta_i,\varepsilon_{-1})=0,\ i=3,\cdots,n.\end{aligned} \tag{18}$$

容易从(16)式看出, $\varepsilon_1,\varepsilon_{-1},\beta_3',\cdots,\beta_n'$ 仍是 V 的一个基,令

$$W=\langle\beta_3',\cdots,\beta_n'\rangle,$$

则

$$V=\langle\varepsilon_1,\varepsilon_{-1}\rangle\oplus W.$$

把 f 看成 W 上的双线性函数,它仍是斜对称的.据归纳假设,存在 W 的一个基

$$\varepsilon_2,\varepsilon_{-2},\cdots,\varepsilon_r,\varepsilon_{-r},\eta_1,\cdots,\eta_s,$$

使得 f 在这个基下的度量矩阵为

$$B_1=\mathrm{diag}\left(\begin{pmatrix}0&1\\-1&0\end{pmatrix},\cdots,\begin{pmatrix}0&1\\-1&0\end{pmatrix},0,\cdots,0\right). \tag{19}$$

从(17)和(18)式得出, $\forall\beta\in W$,有 $f(\varepsilon_1,\beta)=0$ 且 $f(\varepsilon_{-1},\beta)=0$.从而 f 在 V 的一个基

$$\varepsilon_1,\varepsilon_{-1},\varepsilon_2,\varepsilon_{-2},\cdots,\varepsilon_r,\varepsilon_{-r},\eta_1,\cdots,\eta_s$$

下的度量矩阵为

$$B=\mathrm{diag}\left(\begin{pmatrix}0&1\\-1&0\end{pmatrix},\begin{pmatrix}0&1\\-1&0\end{pmatrix},\cdots,\begin{pmatrix}0&1\\-1&0\end{pmatrix},0,\cdots,0\right).$$

由归纳法原理,对一切自然数 n,命题成立. x

从定理 5 得出,如果 V 上的斜对称双线性函数 f 是非退化的,则存在 V 的一个基,使得 f 在此基下的度量矩阵为

$$\operatorname{diag}\left\{\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}, \cdots, \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\right\}. \tag{20}$$

从而得出,$\dim V$ 一定是偶数.

习 题 10.1

1. 在 K^4 中,设 $\alpha=(x_1,x_2,x_3,x_4)'$,$\beta=(y_1,y_2,y_3,y_4)'$,令

$$f(\alpha,\beta)=x_1y_1+x_2y_2+x_3y_3-x_4y_4,$$

(1) 证明:f 是 K^4 上的一个双线性函数;

(2) 求 f 在标准基 $\varepsilon_1,\varepsilon_2,\varepsilon_3,\varepsilon_4$ 下的度量矩阵;

(3) 说明 f 是非退化的;

(4) 说明 f 是对称的;

(5) 求一个向量 $\alpha\neq 0$,使得 $f(\alpha,\alpha)=0$.

2. 证明:当 V 是有限维线性空间时,V 上的双线性函数 f 的左根与右根的维数相同,都等于 $\dim V-\operatorname{rank}_m f$.

3. 设 f 是域 F 上 n 维线性空间 V 上的一个双线性函数.

证明:(1) 映射 $L_f:\alpha\longmapsto\alpha_L$ 是 V 到 V^* 的一个线性映射;映射 $R_f:\beta\longmapsto\beta_R$ 是 V 到 V^* 的一个线性映射;

(2) $\operatorname{Ker}L_f=\operatorname{rad}_L V$,$\operatorname{Ker}R_f=\operatorname{rad}_R V$;

(3) $\operatorname{rank}L_f=\operatorname{rank}_m f=\operatorname{rank}R_f$;

(4) f 是非退化的充分必要条件为 L_f(或 R_f)是线性空间 V 到 V^* 的同构映射.

4. 证明:$M_n(F)$ 上的双线性函数 $f(A,B)=\operatorname{tr}(AB)$ 是非退化的.

5. 设 V 是复数域上的 n 维线性空间,$n\geq 2$.设 f 是 V 上的一个对称双线性函数.

(1) 证明:V 中有非零向量 ξ,使 $f(\xi,\xi)=0$;

(2) 如果 f 是非退化的,则必有线性无关的向量 ξ,η,满足:$f(\xi,\eta)=1$,$f(\xi,\xi)=f(\eta,\eta)=0$.

6. 设 f 是特征不为 2 的域 F 上线性空间 V 的双线性函数,证明:f 是斜对称的充分必要条件为:对任意 $\alpha\in V$,都有 $f(\alpha,\alpha)=0$.

7. 设 V 是域 F 上的线性空间,f 是 V 上的对称的或斜对称的双线性函数.设 W 是 V 的一个子空间,令

$$W^{\perp}\overset{\text{def}}{=}\{\alpha\in V\mid f(\alpha,\beta)=0,\ \forall\,\beta\in W\}.$$

证明:$W^{\perp}$ 是 V 的子空间,称 $W^{\perp}$ 是 W 的正交补.

8. 设 V 是域 F 上的 n 维线性空间,f 是 V 上的非退化的对称双线性函数,W 是 V 的一个子空间.证明:

(1) $\dim W+\dim W^{\perp}=\dim V$;

(2) $(W^{\perp})^{\perp}=W$.

§2　欧几里得空间

我们已经知道,R^n 上的一个内积(α,β)是 R^n 上的一个二元实值函数,且具有对称性,线性,正定性等性质,因此(α,β)是一个正定的对称双线性函数.由此受到启发,我们引出下述概念:

定义1　设 V 是实数域 R 上的一个线性空间,V 上的一个正定对称双线性函数 f 称为 V 上的一个内积(f 是正定的意思是,$\forall\ \alpha\in V$,都有 $f(\alpha,\alpha)\geq 0$,等号成立当且仅当 $\alpha=0$).

习惯上把内积 f 在有序向量对(α,β)上的函数值 $f(\alpha,\beta)$简记成(α,β).

命题1　设 V 是 R 上的一个 n 维线性空间,V 上的双线性函数 f 是正定对称的充分必要条件是,f 在 V 的一个基下的度量矩阵是正定对称的.

证明　V 中取一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$,设 f 在此基下的度量矩阵是 A;α,β 在此基下的坐标分别是 X,Y,则 $f(\alpha,\beta)=X'AY$.

f 是对称的双线性函数当且仅当 A 是对称矩阵.

f 是正定的　　$\forall\ \alpha\in V$,且 $\alpha\neq 0$,有 $f(\alpha,\alpha)>0$

　　　　　　　$\forall\ X\in R^n$,且 $X\neq 0$,有 $X'AX>0$

　　　　　　　A 是正定矩阵.　　　　　　　　　　x

例1　R^3 中,设 $\alpha=(a_1,a_2,a_3)$,$\beta=(b_1,b_2,b_3)$,规定

$$(\alpha,\beta)\stackrel{def}{=}a_1b_1+2a_2b_2+3a_3b_3, \tag{1}$$

容易验证,(α,β)是 R^3 上的一个内积,它与 R^3 上的标准内积不同.

例2　在实数域 R 上的线性空间 $M_n(R)$中,规定一个二元函数为

$$(A,B)\stackrel{def}{=}\mathrm{tr}(AB'), \tag{2}$$

容易验证,(A,B)是 $M_n(R)$上的一个内积.

例3　在 $C[a,b]$中,规定一个二元函数为

$$(f,g)\stackrel{def}{=}\int_a^b f(x)g(x)dx, \tag{3}$$

容易验证,(f,g)是 $C[a,b]$上的一个内积.

定义2　实数域 R 上的线性空间 V 如果给定了一个内积,则称 V 是一个实内积空间.有限维的实内积空间 V 称为欧几里得空间,此时把线性空间 V 的维数叫做欧几里得空间 V 的维数.

在实内积空间 V 中,由于有了内积的概念,因此就会有长度、角度、正交、距离等度量概念.

定义 3　非负实数　(α,α)称为向量 α 的长度,记作$|\alpha|$(或者$\|\alpha\|$).

根据内积的正定性,零向量的长度为 0,非零向量的长度是正数.我们有

$$|k\alpha| = |k||\alpha|, \quad " \alpha \in V, k \in R.$$

证明　$|k\alpha| = \quad (k\alpha, k\alpha) = \quad k^2(\alpha,\alpha) = |k||\alpha|.$　　x

长度为 1 的向量称为单位向量.如果 $\alpha \neq 0$,则$\frac{1}{|\alpha|}\alpha$是一个单位向量.把 α 变成$\frac{1}{|\alpha|}\alpha$称为把 α 单位化.

定理 2(Cauchy - Buniakowski 不等式)　在实内积空间 V 中,对于任意向量 α,β,有

$$|(\alpha,\beta)| \leq |\alpha||\beta|, \tag{4}$$

等号成立当且仅当 α,β 线性相关.

证明　如果 α,β 线性相关,则 $\alpha = 0$ 或者 $\beta = k\alpha$.如果 $\alpha = 0$,则

$$|(0,\beta)| = |0(0,\beta)| = 0 = |0||\beta|;$$

如果 $\beta = k\alpha$,则

$$|(\alpha,\beta)| = |(\alpha,k\alpha)| = |k(\alpha,\alpha)| = |k||\alpha|^2 = |\alpha||k\alpha| = |\alpha||\beta|.$$

如果 α,β 线性无关,则对一切实数 t,有 $\beta \neq t\alpha$,从而有 $t\alpha - \beta \neq 0$.根据内积的正定性得,$" t \in R$,有

$$0 < (t\alpha - \beta, t\alpha - \beta) = t^2|\alpha|^2 - 2t(\alpha,\beta) + |\beta|^2, \tag{5}$$

于是(5)式右端的 t 的 2 次多项式的判别式小于 0,即

$$4(\alpha,\beta)^2 - 4|\alpha|^2|\beta|^2 < 0.$$

由此得出

$$|(\alpha,\beta)| < |\alpha||\beta|.$$　　x

定义 4　实内积空间中,两个非零向量 α,β 的夹角$\langle\alpha,\beta\rangle$规定为

$$\langle\alpha,\beta\rangle \overset{\text{def}}{} \arccos\frac{(\alpha,\beta)}{|\alpha||\beta|}, \tag{6}$$

于是 $0 \leq \langle\alpha,\beta\rangle \leq \pi$.

从(6)式得出

$$\langle\alpha,\beta\rangle = \frac{\pi}{2} \qquad (\alpha,\beta) = 0,$$

于是,有

定义 5　如果$(\alpha,\beta) = 0$,则称 α 与 β 正交,记为 $\alpha \perp \beta$.

推论 3　在实内积空间 V 中,三角形不等式成立,即对于任意 $\alpha,\beta \in V$,有

$$|\alpha + \beta| \leq |\alpha| + |\beta|. \tag{7}$$

证明　$|\alpha + \beta|^2 = (\alpha + \beta, \alpha + \beta) = |\alpha|^2 + 2(\alpha,\beta) + |\beta|^2$

$$\leq |\alpha|^2 + 2|\alpha||\beta| + |\beta|^2 = (|\alpha| + |\beta|)^2.$$

由此得出
$$|\alpha+\beta|\leq|\alpha|+|\beta|.$$
x

推论4　在实内积空间 V 中,勾股定理成立,即如果 α 与 β 正交,则
$$|\alpha+\beta|^2=|\alpha|^2+|\beta|^2. \tag{8}$$

证明
$$|\alpha+\beta|^2=(\alpha+\beta,\alpha+\beta)=|\alpha|^2+|\beta|^2.$$
x

定义6　在实内积空间 V 中,对于任意 $\alpha,\beta\in V$,规定
$$d(\alpha,\beta)\overset{\text{def}}{=}|\alpha-\beta|, \tag{9}$$
称 $d(\alpha,\beta)$ 是 α 与 β 的距离.

容易验证,对任意 $\alpha,\beta,\gamma\in V$,有

1°　$d(\alpha,\beta)=d(\beta,\alpha)$　(对称性);

2°　$d(\alpha,\beta)\geq 0$,等号成立当且仅当 $\alpha=\beta$　(正定性);

3°　$d(\alpha,\gamma)\leq d(\alpha,\beta)+d(\beta,\gamma)$　(三角形不等式).

在欧几里得空间 V 中,由两两正交的非零向量组成的向量组称为正交向量组.由两两正交的单位向量组成的向量组称为正交单位向量组.

命题5　在欧几里得空间 V 中,正交向量组一定线性无关.

证明　与第4章§6的命题2的证明一样. x

在 n 维欧几里得空间 V 中, n 个向量组成的正交向量组一定是 V 的一个基,称它为正交基; n 个单位向量组成的正交向量组称为 V 的一个标准正交基(或规范正交基).

与第4章§6的定理4的证明方法完全一样,可以证明下述结论:

定理6　设 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 是欧几里得空间 V 的一个线性无关的向量组,令
$$\beta_1=\alpha_1,$$
$$\beta_2=\alpha_2-\frac{(\alpha_2,\beta_1)}{(\beta_1,\beta_1)}\beta_1,$$
$$\cdots\cdots\cdots\cdots$$
$$\beta_s=\alpha_s-\sum_{j=1}^{s-1}\frac{(\alpha_s,\beta_j)}{(\beta_j,\beta_j)}\beta_j, \tag{10}$$
则 $\beta_1,\beta_2,\cdots,\beta_s$ 是正交向量组,并且 $\beta_1,\beta_2,\cdots,\beta_s$ 与 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 等价. x

定理6中把线性无关的向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 变成与它等价的正交向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 的过程称为施密特(Schimidt)正交化.只要再将每个 β_j 单位化,就可得到一个与 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 等价的正交单位向量组 $\eta_1,\eta_2,\cdots,\eta_s$.因此把 n 维欧几里得空间 V 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 经过施密特正交化过程,以及单位化,就可得到 V 的一个标准正交基 $\eta_1,\eta_2,\cdots,\eta_n$.

由定义知道, n 维欧几里得空间 V 中,向量组 $\eta_1,\eta_2,\cdots,\eta_n$ 是 V 的一个标准正交基当且仅当

$$(\eta_i,\eta_j)=\delta_{ij},\quad i,j=1,2,\cdots,n. \tag{11}$$

利用标准正交基容易计算向量的内积.设 α,β 在 V 的标准正交基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的坐标分别是 $X=(x_1,x_2,\cdots,x_n)'$, $Y=(y_1,y_2,\cdots,y_n)'$,则

$$(\alpha,\beta)=\left(\sum_{i=1}^{n}x_i\eta_i,\sum_{j=1}^{n}y_j\eta_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}x_iy_j(\eta_i,\eta_j)$$

$$=\sum_{i=1}^{n}x_iy_i=X'Y. \tag{12}$$

利用标准正交基,向量的坐标的分量可以用内积表达.设 α 在标准正交基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的坐标为 $(x_1,x_2,\cdots,x_n)'$,则

$$\alpha=\sum_{i=1}^{n}x_i\eta_i.$$

两边用 η_j 作内积,得

$$(\alpha,\eta_j)=\left(\sum_{i=1}^{n}x_i\eta_i,\eta_j\right)=\sum_{i=1}^{n}x_i(\eta_i,\eta_j)=x_j,$$

因此

$$\alpha=\sum_{i=1}^{n}(\alpha,\eta_i)\eta_i. \tag{13}$$

(13)式称为 α 的傅里叶(Fourier)展开,其中每个系数 (α,η_i) 都称为 α 的傅里叶(Fourier)系数.

命题 7 欧几里得空间 V 中,标准正交基到标准正交基的过渡矩阵是正交矩阵.

证明 设 $\eta_1,\eta_2,\cdots,\eta_n$ 与 $\beta_1,\beta_2,\cdots,\beta_n$ 是 V 的两个标准正交基,P 是基 $\eta_1,\eta_2,\cdots,\eta_n$ 到基 $\beta_1,\beta_2,\cdots,\beta_n$ 的过渡矩阵,即

$$(\beta_1,\beta_2,\cdots,\beta_n)=(\eta_1,\eta_2,\cdots,\eta_n)P, \tag{14}$$

于是 P 的第 j 列 Y_j 是 β_j 在标准正交基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的坐标,$j=1,2,\cdots,n$,从而

$$(\beta_i,\beta_j)=Y_i'Y_j,\quad i,j=1,2,\cdots,n. \tag{15}$$

由于 $\beta_1,\beta_2,\cdots,\beta_n$ 是标准正交基,因此

$$(\beta_i,\beta_j)=\delta_{ij},\quad i,j=1,2,\cdots,n. \tag{16}$$

从(15)式和(16)式得

$$Y_i'Y_j=\delta_{ij},\quad i,j=1,2,\cdots,n. \tag{17}$$

据第 4 章§6 的定理 1 得,P 是正交矩阵. x

命题 8 n 维欧几里得空间 V 中,设 $\eta_1,\eta_2,\cdots,\eta_n$ 是 V 的一个标准正交基,向量组 $\beta_1,\beta_2,\cdots,\beta_n$ 满足

$$(\beta_1, \beta_2, \cdots, \beta_n) = (\eta_1, \eta_2, \cdots, \eta_n) P,$$

其中 P 是正交矩阵,则 $\beta_1, \beta_2, \cdots, \beta_n$ 是 V 的一个标准正交基.

证明　P 的第 j 列 Y_j 是 β_j 在标准正交基 $\eta_1, \eta_2, \cdots, \eta_n$ 下的坐标, $j = 1, 2, \cdots, n$. 由于 P 是正交矩阵,因此

$$Y_i' Y_j = \delta_{ij}, \quad i, j = 1, 2, \cdots, n, \tag{18}$$

于是

$$(\beta_i, \beta_j) = Y_i' Y_j = \delta_{ij}, \quad i, j = 1, 2, \cdots, n, \tag{19}$$

因此 $\beta_1, \beta_2, \cdots, \beta_n$ 是 V 的一个标准正交基.　x

对于实数域上的一个线性空间 V,当指定不同的内积时, V 便成为不同的实内积空间,这些实内积空间之间有什么关系？不同的实线性空间,各自指定了一个内积,成为实内积空间后,它们之间又有什么关系？

定义 7　设 V 和 V' 都是实内积空间,如果存在 V 到 V' 的一个双射 σ,使得对于任意 $\alpha, \beta \in V$, $k \in R$,有

$$\begin{aligned} &\sigma(\alpha + \beta) = \sigma(\alpha) + \sigma(\beta), \\ &\sigma(k\alpha) = k\sigma(\alpha), \\ &(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta), \end{aligned}$$

则称 σ 是实内积空间 V 到 V' 的一个同构映射,此时称 V 与 V' 是同构的,记作 $V \cong V'$.

从定义 7 看出,实内积空间 V 到 V' 的一个同构映射 σ 首先是实线性空间 V 到 V' 的一个同构映射,其次 σ 还保持内积,因此 σ 既具有线性空间的同构映射的性质,又还具有与内积有关的性质,譬如 σ 把 V 的一个标准正交基映成 V' 的一个标准正交基. 理由如下: 设 $\eta_1, \eta_2, \cdots, \eta_n$ 是欧几里得空间 V 的一个标准正交基,则

$$(\eta_i, \eta_j) = \delta_{ij}, \quad i, j = 1, 2, \cdots, n.$$

从而

$$(\sigma(\eta_i), \sigma(\eta_j)) = (\eta_i, \eta_j) = \delta_{ij}, \quad i, j = 1, 2, \cdots, n.$$

因此 $\sigma(\eta_1), \sigma(\eta_2), \cdots, \sigma(\eta_n)$ 不仅是 V' 的一个基(据线性空间的同构映射的性质),而且是 V' 的一个标准正交基.

定理 9　两个欧几里得空间同构的充分必要条件是它们的维数相同.

证明　设 V 与 V' 都是欧几里得空间.

必要性. 由于 V 与 V' 作为线性空间也同构,因此它们的维数相同.

充分性. 在 V 与 V' 中各取一个标准正交基: $\eta_1, \eta_2, \cdots, \eta_n$ 与 $\delta_1, \delta_2, \cdots, \delta_n$. 令

$$\sigma: \quad V \to V'$$
$$\alpha=\sum_{i=1}^{n} x_i\eta_i \longmapsto \sum_{i=1}^{n} x_i\delta_i,$$

则 σ 是线性空间 V 到 V' 的一个同构映射. 设 $\beta=\sum_{i=1}^{n} y_i\eta_i$, 则 $\sigma(\beta)=\sum_{i=1}^{n} y_i\delta_i$. 又有 $\sigma(\alpha)=\sum_{i=1}^{n} x_i\delta_i$, 于是

$$(\sigma(\alpha),\sigma(\beta))=\sum_{i=1}^{n} x_iy_i=(\alpha,\beta).$$

因此 σ 是实内积空间 V 到 V' 的一个同构映射, 从而 $V\cong V'$. ■

从定理 8 得出, 任一 n 维欧几里得空间 V 都与装备了标准内积的欧几里得空间 R^n 同构, 并且一个同构映射是:

$$\sigma: \quad V \to R^n$$
$$\alpha=\sum_{i=1}^{n} x_i\eta_i \longmapsto (x_1, x_2, \cdots, x_n)',$$

其中 $\eta_1, \eta_2, \cdots, \eta_n$ 是 V 的一个标准正交基.

可以证明, 同构作为实内积空间之间的关系具有反身性、对称性、传递性.

习 题 10.2

1. 在 R^2 中, 对于任意 $\alpha=(x_1, x_2)$, $\beta=(y_1, y_2)$, 规定

$$(\alpha,\beta)\overset{\text{def}}{=} x_1y_1-x_1y_2-x_2y_1+4x_2y_2,$$

判断 (α,β) 是不是 R^2 上的一个内积.

2. 在实线性空间 $M_n(R)$ 中, 规定一个二元函数为

$$f(A,B)\overset{\text{def}}{=}\operatorname{tr}(AB),$$

判断 $f(A,B)$ 是不是 $M_n(R)$ 上的一个内积.

3. 在 R^n 中, 对于任意两个列向量 α,β, 规定

$$(\alpha,\beta)\overset{\text{def}}{=}\alpha'A\beta,$$

其中 A 是一个 n 级正定矩阵. 证明: (α,β) 是 R^n 上的一个内积.

4. 在 R^4 中, 给定了标准内积, 求 α 与 β 的夹角 $\langle\alpha,\beta\rangle$, 其中

$$\alpha=(1,-1,4,0),\quad \beta=(3,1,-2,2).$$

5. 在 $R[x]_3$ 中, 给定一个内积为

$$(f(x),g(x))\overset{\text{def}}{=}\int_{-1}^{1} f(x)g(x)\,dx,$$

求 $R[x]_3$ 的一个标准正交基.

6. 设 V 是3维欧几里得空间,$\alpha_1,\alpha_2,\alpha_3$ 是 V 的一个基.令

$$A=\begin{pmatrix}(\alpha_1,\alpha_1) & (\alpha_1,\alpha_2) & (\alpha_1,\alpha_3)\\ (\alpha_2,\alpha_1) & (\alpha_2,\alpha_2) & (\alpha_2,\alpha_3)\\ (\alpha_3,\alpha_1) & (\alpha_3,\alpha_2) & (\alpha_3,\alpha_3)\end{pmatrix},$$

称 A 是基 $\alpha_1,\alpha_2,\alpha_3$ 的度量矩阵.设

$$A=\begin{pmatrix}1 & 0 & 1\\ 0 & 10 & -2\\ 1 & -2 & 2\end{pmatrix},$$

求 V 的一个标准正交基.

7. 设 η_1,η_2,η_3 是3维欧几里得空间 V 的一个标准正交基,令

$$\beta_1=\frac{1}{3}(2\eta_1-\eta_2+2\eta_3),$$

$$\beta_2=\frac{1}{3}(2\eta_1+2\eta_2-\eta_3),$$

$$\beta_3=\frac{1}{3}(\eta_1-2\eta_2-2\eta_3).$$

证明:β_1,β_2,β_3 是 V 的一个标准正交基.

8. 设 $\eta_1,\eta_2,\eta_3,\eta_4,\eta_5$ 是5维欧几里得空间 V 的一个标准正交基,$V_1=\langle\alpha_1,\alpha_2,\alpha_3\rangle$,其中

$$\alpha_1=\eta_1+2\eta_3-\eta_5,$$

$$\alpha_2=\eta_2-\eta_3+\eta_4,$$

$$\alpha_3=-\eta_2+\eta_3+\eta_5.$$

(1) 求(α_i,α_j),$1\leq i,j\leq 3$;

(2) 求 V_1 的一个正交基.

*9. 在 R^2 中指定一个内积为

$$(\alpha,\beta)\overset{\text{def}}{=}x_1y_1+2x_2y_2,$$

其中 $\alpha=(x_1,x_2)'$,$\beta=(y_1,y_2)'$,把这个欧几里得空间记作 V.找出 V 到指定标准内积的欧几里得空间 R^2 的一个同构映射.

§3　正交补,正交投影

几何空间中,设 U 是过原点 O 的一个平面,l 是过原点 O 且与平面 U 垂直的直线,则 l 的每一个向量与平面 U 的每一个向量都正交,我们称 l 是 U 的正交补.

一般的欧几里得空间 V 中,它的任一子空间 U 有没有正交补的概念?它有什么性质?本节就来讨论这一问题.

设 V 是一个实内积空间，V_1 是 V 的任一线性子空间．显然 V_1 中的每两个向量有内积(按照 V 上指定的内积进行计算)，因此 V_1 也成为实内积空间，称 V_1 是实内积空间 V 的一个子空间．

定义 1　设 V 是实内积空间，S 是 V 的一个非空子集．我们把 V 中与 S 的每一个向量都正交的所有向量组成的集合叫做 S 的正交补，记作 $S^{\perp}$．即

$$S^{\perp} \overset{\text{def}}{=} \{\alpha \in V \mid (\alpha, \beta) = 0, \forall \beta \in S\}. \tag{1}$$

容易验证，$S^{\perp}$ 是 V 的一个子空间．

如果实内积空间 V 有一个有限维子空间 U，则 V 的结构可以由 U 与 $U^{\perp}$ 的直和来刻画，即我们有下述重要结论：

定理 1　设 U 是实内积空间 V 的一个有限维子空间，则

$$V = U \oplus U^{\perp}. \tag{2}$$

证明　先证 $V = U + U^{\perp}$．在 U 中取一个标准正交基 $\varepsilon_1, \cdots, \varepsilon_m$．任取 $\alpha \in V$，令

$$\alpha_1 = (\alpha, \varepsilon_1)\varepsilon_1 + \cdots + (\alpha, \varepsilon_m)\varepsilon_m \in U, \tag{3}$$

$$\alpha_2 = \alpha - \alpha_1. \tag{4}$$

因为对于 $j = 1, \cdots, m$，有

$$\begin{aligned}(\alpha_2, \varepsilon_j) &= (\alpha - \alpha_1, \varepsilon_j) = (\alpha, \varepsilon_j) - \left(\sum_{i=1}^{m} (\alpha, \varepsilon_i)\varepsilon_i, \varepsilon_j\right) \\ &= (\alpha, \varepsilon_j) - (\alpha, \varepsilon_j) = 0,\end{aligned}$$

所以 $\alpha_2 \in U^{\perp}$．从 $\alpha = \alpha_1 + \alpha_2$ 知道 $V = U + U^{\perp}$．

再证 $U \cap U^{\perp} = 0$．假设 $\gamma \in U \cap U^{\perp}$，则 $(\gamma, \gamma) = 0$，从而得 $\gamma = 0$．

综上述得，$V = U \oplus U^{\perp}$．　x

例如，在几何空间 V 中，设 U 是过原点的一个平面，则 $U^{\perp}$ 是过原点并且与平面 U 垂直的直线．显然，$V = U \oplus U^{\perp}$．

从定理 1 知道，如果 U 是实内积空间 V 的一个有限维子空间，则 $V = U \oplus U^{\perp}$．从而 V 中每个向量 α 能唯一地表示成

$$\alpha = \alpha_1 + \alpha_2, \quad \alpha_1 \in U, \quad \alpha_2 \in U^{\perp}. \tag{5}$$

从而有 V 上的线性变换 $P_U : \alpha \longmapsto \alpha_1$．我们把 P_U 称为 V 在 U 上的正交投影，把 α_1 称为 α 在 U 上的正交投影．从(5)式得出，$\alpha_1 \in U$ 是 α 在 U 上的正交投影当且仅当 $\alpha - \alpha_1 \in U^{\perp}$．

定理 2　设 U 是欧几里得空间 V 的一个子空间，对于 $\alpha \in V$，$\alpha_1 \in U$ 是 α 在 U 上的正交投影当且仅当

$$d(\alpha, \alpha_1) \leq d(\alpha, \gamma), \quad \forall \gamma \in U. \tag{6}$$

证明 必要性.设 $\alpha_1\in U$ 是 α 在 U 上的正交投影,则 $\alpha-\alpha_1\in U^{\perp}$.从而 $\forall\ \gamma\in U$,有

$$(\alpha-\alpha_1)\perp(\alpha_1-\gamma).$$

由勾股定理得

$$\begin{aligned}|\alpha-\alpha_1|^2+|\alpha_1-\gamma|^2&=|(\alpha-\alpha_1)+(\alpha_1-\gamma)|^2\\&=|\alpha-\gamma|^2.\end{aligned}$$

由此得出,$|\alpha-\alpha_1|\leq|\alpha-\gamma|$,即 $d(\alpha,\alpha_1)\leq d(\alpha,\gamma)$.

充分性.设(6)式成立.假设 δ 是 α 在 U 上的正交投影,据必要性得,$d(\alpha,\delta)\leq d(\alpha,\alpha_1)$.结合(6)式得

$$d(\alpha,\delta)=d(\alpha,\alpha_1).$$

由于 $(\alpha-\delta)\in U^{\perp}$,$(\delta-\alpha_1)\in U$,因此

$$|\alpha-\alpha_1|^2=|(\alpha-\delta)+(\delta-\alpha_1)|^2=|\alpha-\delta|^2+|\delta-\alpha_1|^2.$$

由此得出,$|\delta-\alpha_1|^2=0$.因此 $\delta=\alpha_1$. x

正交投影有重要的应用.

实际问题中从观测数据列出的线性方程组 $AX=\beta$ 可能无解,其中,$A=(a_{ij})_{s\times n}$,$\beta=(b_1,\cdots,b_s)'$.把 A 的行向量组记作 $\gamma_1,\gamma_2,\cdots,\gamma_s$.当 $AX=\beta$ 无解时,由于实际问题的需要,我们想找一个列向量 $\alpha=(c_1,c_2,\cdots,c_n)'$,使得当 $x_1=c_1,x_2=c_2,\cdots,x_n=c_n$ 时,下述式子

$$\sum_{i=1}^{s}[(a_{i1}c_1+a_{i2}c_2+\cdots+a_{in}c_n)-b_i]^2 \tag{7}$$

达到最小值,这个列向量 α 称为线性方程组的最小二乘解.如何求 $AX=\beta$ 的最小二乘解?

(7)式是平方和的形式,这使人联想到它是欧几里得空间 R^s(指定的内积是标准内积)中某个向量的长度的平方.这个向量的第 i 个分量是

$$(a_{i1}c_1+a_{i2}c_2+\cdots+a_{in}c_n)-b_i=\gamma_i\alpha-b_i,\quad i=1,2,\cdots,s,$$

因此这个向量是

$$\begin{pmatrix}\gamma_1\alpha-b_1\\\gamma_2\alpha-b_2\\\cdots\\\gamma_s\alpha-b_s\end{pmatrix}=\begin{pmatrix}\gamma_1\alpha\\\gamma_2\alpha\\\cdots\\\gamma_s\alpha\end{pmatrix}-\begin{pmatrix}b_1\\b_2\\\cdots\\b_s\end{pmatrix}=A\alpha-\beta, \tag{8}$$

于是 α 是 $AX=\beta$ 的最小二乘解当且仅当 $A\alpha-\beta$ 的长度 $|A\alpha-\beta|$ 最小,即对任意 $X\in R^n$,有 $|A\alpha-\beta|\leq|AX-\beta|$.

设 A 的列向量组是 $\alpha_1,\alpha_2,\cdots,\alpha_n$,则 $X\in R^n$ 当且仅当

$$AX=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n\in\langle\alpha_1,\alpha_2,\cdots,\alpha_n\rangle.$$

把 A 的列空间 $\langle \alpha_1, \alpha_2, \cdots, \alpha_n \rangle$ 记作 U,于是

$$
\begin{aligned}
&\alpha \text{ 是 } AX = \beta \text{ 的最小二乘解} \\
&|A\alpha - \beta| \leq |AX - \beta|, \forall X \in R^n \\
&|A\alpha - \beta| \leq |\gamma - \beta|, \forall \gamma \in U \\
&d(A\alpha, \beta) \leq d(\gamma, \beta), \forall \gamma \in U \\
&A\alpha \text{ 是 } \beta \text{ 在 } U \text{ 上的正交投影} \\
&\beta - A\alpha \in U^{\perp} \\
&(\beta - A\alpha, \alpha_j) = 0, j = 1, 2, \cdots, n \\
&\alpha_j'(\beta - A\alpha) = 0, j = 1, 2, \cdots, n \\
&A'(\beta - A\alpha) = 0 \\
&A'A\alpha = A\beta \\
&\alpha \text{ 是}(A'A)X = A\beta \text{ 的解.}
\end{aligned}
$$

由于

$$\operatorname{rank}(A'A, A\beta) = \operatorname{rank}(A'(A, B)) \leq \operatorname{rank}(A') = \operatorname{rank}(A'A),$$

$$\operatorname{rank}(A'A, A\beta) \geq \operatorname{rank}(A'A),$$

因此

$$\operatorname{rank}(A'A, A\beta) = \operatorname{rank}(A'A),$$

从而线性方程$(A'A)X = A\beta$一定有解.这样我们把求线性方程组 $AX = \beta$ 的最小二乘解的问题归结为求线性方程组$(A'A)X = A\beta$的解.

习 题 10.3

1. 设 U 是欧几里得空间 R^4(指定标准内积)的一个子空间,$U = \langle \alpha_1, \alpha_2 \rangle$,其中

$$\alpha_1 = (1,1,2,1), \quad \alpha_2 = (1,0,0,-2).$$

求 $U^{\perp}$ 的维数和一个正交基.

2. 设 V 是一个 n 维欧几里得空间,$\alpha \in V$ 且 $\alpha \neq 0$,求 $\langle \alpha \rangle^{\perp}$ 的维数.

3. 设 U 是 n 维欧几里得空间 V 的一个子空间,证明:

$$(U^{\perp})^{\perp} = U.$$

4. 证明:欧几里得空间 R^n(指定标准内积)的任一子空间 U 是一个齐次线性方程组的解空间.

5. 设 U 是 n 维欧几里得空间 V 的一个子空间,在 U 中取一个标准正交基 $\eta_1, \eta_2, \cdots, \eta_m$,则 α 在 U 上的正交投影 α_1 为

$$\alpha_1 = \sum_{i=1}^{m} (\alpha, \eta_i) \eta_i.$$

6. 在欧几里得空间 R^3(指定标准内积)中,设 $U = \langle \gamma_1, \gamma_2 \rangle$,其中 $\gamma_1 = (1,2,1)$,$\gamma_2 = (1,

0, -2). 求 $\alpha=(1,-3,0)$ 在 U 上的正交投影 α_1.

7. 设 U 是欧几里得空间 V 的一个子空间,则 V 在 U 上的正交投影 P 具有下述性质:

$$(P\alpha,\beta)=(\alpha,P\beta),\quad \forall\,\alpha,\beta\in V.$$

8. $\{\varepsilon_1,\cdots,\varepsilon_m\}$ 是实内积空间 V 的一个正交单位向量组. 证明:对于一切 $\alpha\in V$,有

$$\sum_{i=1}^{m}(\alpha,\varepsilon_i)^2\le|\alpha|^2,$$

等号成立当且仅当 $\alpha=\sum_{i=1}^{m}(\alpha,\varepsilon_i)\varepsilon_i$. 这个不等式称为 Bessel 不等式.

9. 在欧氏空间 $R[x]_4$ 中,其指定的内积为

$$(f,g)=\int_0^1 f(x)g(x)\mathrm{d}x,$$

设 W 是由零次多项式和零多项式组成的子空间,求 $W^{\perp}$ 以及它的一个基.

10. 实内积空间 $M_n(R)$,指定的内积为 $(A,B)=\mathrm{tr}(AB')$. 设 W 是由所有对角矩阵组成的子空间,求 $W^{\perp}$ 以及 $W^{\perp}$ 的一个标准正交基.

§4　正交变换与对称变换

平面到自身上的一个变换 σ 如果保持两点间的距离不变,则称 σ 是平面上的正交变换.

在实内积空间中也有正交变换的概念. 本节来讨论正交变换.

定义 1　实内积空间 V 到自身的满射 $\mathbf{A}$ 如果保持向量的内积不变,即

$$(\mathbf{A}\alpha,\mathbf{A}\beta)=(\alpha,\beta),\quad \forall\,\alpha,\beta\in V,\tag{1}$$

则称 $\mathbf{A}$ 是 V 上的一个正交变换.

从定义 1 容易看出, V 上的正交变换保持向量的长度不变.

命题 1　实内积空间 V 上的正交变换 $\mathbf{A}$ 一定是线性变换.

证明　先证 $\mathbf{A}(\alpha+\beta)=\mathbf{A}\alpha+\mathbf{A}\beta$. 由于

$$\begin{aligned}&|\mathbf{A}(\alpha+\beta)-(\mathbf{A}\alpha+\mathbf{A}\beta)|^2\\=&|\mathbf{A}(\alpha+\beta)|^2-2(\mathbf{A}(\alpha+\beta),\mathbf{A}\alpha+\mathbf{A}\beta)+|\mathbf{A}\alpha+\mathbf{A}\beta|^2\\=&|\alpha+\beta|^2-2(\mathbf{A}(\alpha+\beta),\mathbf{A}\alpha)-2(\mathbf{A}(\alpha+\beta),\mathbf{A}\beta)\\&+|\mathbf{A}\alpha|^2+2(\mathbf{A}\alpha,\mathbf{A}\beta)+|\mathbf{A}\beta|^2\\=&|\alpha+\beta|^2-2(\alpha+\beta,\alpha)-2(\alpha+\beta,\beta)+|\alpha|^2+2(\alpha,\beta)+|\beta|^2\\=&|\alpha+\beta|^2-2(\alpha+\beta,\alpha+\beta)+|\alpha+\beta|^2=0,\end{aligned}$$

因此 $\mathbf{A}(\alpha+\beta)-(\mathbf{A}\alpha+\mathbf{A}\beta)=0$,即 $\mathbf{A}(\alpha+\beta)=\mathbf{A}\alpha+\mathbf{A}\beta$.

同理可证, $\mathbf{A}(k\alpha)=k\mathbf{A}\alpha,\ \forall\,\alpha\in V,k\in R$. 因此 $\mathbf{A}$ 是 V 上的一个线性变换.

X

命题 2 实内积空间 V 上的正交变换 A 一定是单射,从而 A 是可逆的.

证明 设 $A\alpha = A\beta$,则 $A(\alpha - \beta) = 0$.从而

$$|\alpha - \beta| = |A(\alpha - \beta)| = |0| = 0.$$

因此 $\alpha - \beta = 0$,即 $\alpha = \beta$.这证明了 A 是单射.又由定义知,A 是满射.因此 A 是双射,从而 A 是可逆的. x

从命题 1、命题 2 以及定义 1 立即得出

命题 3 实内积空间 V 上的一个变换 A 是正交变换当且仅当 A 是 V 到自身的一个同构映射. x

从实内积空间之间同构关系的对称性和传递性得出,正交变换的逆变换仍是正交变换;正交变换的乘积仍是正交变换.

容易看出,正交变换还保持两个非零向量的夹角不变,保持正交性不变,保持两个向量之间的距离不变.

命题 4 n 维欧几里得空间 V 上的线性变换 A 是正交变换

A 把 V 的标准正交基映成标准正交基

A 在 V 的标准正交基下的矩阵 A 是正交矩阵.

证明 第一个等价条件的必要性.设 A 是 V 上的一个正交变换,则 A 是实内积空间 V 到自身的一个同构映射,从而 A 把 V 的标准正交基映成标准正交基.

充分性.设 V 上的线性变换 A 把 V 的一个标准正交基 $\eta_1, \eta_2, \cdots, \eta_n$ 映成标准正交基 $\gamma_1, \gamma_2, \cdots, \gamma_n$,则 A 是满射.对于任意 $\alpha, \beta \in V$,设 $\alpha = \sum_{i=1}^{n} x_i \eta_i$,$\beta = \sum_{i=1}^{n} y_i \eta_i$,则

$$A\alpha = \sum_{i=1}^{n} x_i A\eta_i = \sum_{i=1}^{n} x_i \gamma_i,$$

$$A\beta = \sum_{i=1}^{n} y_i A\eta_i = \sum_{i=1}^{n} y_i \gamma_i.$$

因此

$$(A\alpha, A\beta) = \sum_{i=1}^{n} x_i y_i = (\alpha, \beta),$$

从而 A 是 V 上的正交变换.

第二个等价条件.设 $\eta_1, \eta_2, \cdots, \eta_n$ 是 V 的一个标准正交基,并且 $A(\eta_1, \eta_2, \cdots, \eta_s) = (\eta_1, \eta_2, \cdots, \eta_n)A$.则据本章§2 命题 7 和命题 8 得

$A\eta_1, A\eta_2, \cdots, A\eta_n$ 是 V 的一个标准正交基

A 是正交矩阵. x

既然欧几里得空间 V 上的正交变换 $\mathcal{A}$ 在 V 的任一标准正交基下的矩阵 A 是正交矩阵,而正交矩阵的行列式等于 1 或者 -1,因此正交变换的行列式等于 1 或 -1.行列式等于 1 的正交变换称为第一类的(或旋转);行列式等于 -1 的正交变换称为第二类的.

在习题 10.3 的第 7 题中,我们指出欧几里得空间 V 在它的子空间 U 上的正交投影 P 具有下述性质:

$$(P\alpha, \beta) = (\alpha, P\beta), \ \forall \alpha, \beta \in V.$$

由此受到启发,我们引出下述概念.

定义 2　实内积空间 V 上的线性变换 $\mathcal{A}$ 如果满足

$$(\mathcal{A}\alpha, \beta) = (\alpha, \mathcal{A}\beta), \ \forall \alpha, \beta \in V,$$

则称 $\mathcal{A}$ 是对称变换.

命题 5　n 维欧几里得空间 V 上的线性变换 $\mathcal{A}$ 是对称变换当且仅当 $\mathcal{A}$ 在 V 的任意一个标准正交基下的矩阵是对称矩阵.

证明　任取 V 的一个标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$.设

$$\mathcal{A}(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n) A.$$

则 $\mathcal{A}\alpha_i$ 在标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的坐标的第 j 个分量为 $a_{ji} = (\mathcal{A}\alpha_i, \alpha_j)$, $i, j = 1, 2, \cdots, n$.因此

$\mathcal{A}$ 是对称变换　$(\mathcal{A}\alpha, \beta) = (\alpha, \mathcal{A}\beta), \ \forall \alpha, \beta \in V$

$(\mathcal{A}\alpha_i, \alpha_j) = (\alpha_i, \mathcal{A}\alpha_j), 1 \leq i, j \leq n$

$a_{ji} = a_{ij}, 1 \leq i, j \leq n$

A 是对称矩阵.　x

命题 6　设 $\mathcal{A}$ 是实内积空间 V 上的一个对称变换,如果 W 是 $\mathcal{A}$ 的不变子空间,则 $W^{\perp}$ 也是 $\mathcal{A}$ 的不变子空间.

证明　任取 $\beta \in W^{\perp}$,要证 $\mathcal{A}\beta \in W^{\perp}$.任取 $\alpha \in W$,有 $\mathcal{A}\alpha \in W$,于是 $(\mathcal{A}\alpha, \beta) = 0$.从而

$$(\alpha, \mathcal{A}\beta) = (\mathcal{A}\alpha, \beta) = 0,$$

因此 $\mathcal{A}\beta \in W^{\perp}$,即 $W^{\perp}$ 是 $\mathcal{A}$ 的不变子空间.　x

我们知道,实对称矩阵一定正交相似于一个对角矩阵,由此得出

定理 7　设 $\mathcal{A}$ 是 n 维欧几里得空间 V 上的一个线性变换,则 $\mathcal{A}$ 是对称变换当且仅当 V 中存在一个标准正交基,使得 $\mathcal{A}$ 在该基下的矩阵为对角矩阵.　x

习　题　10.4

1. 设 V 是实内积空间,证明: V 上的正交变换 $\mathcal{A}$ 如果有特征值,则它的特征值必为 1

或 -1.

2. 设 V 是 n 维欧几里得空间，η 是 V 中一个单位向量，设 P 是 V 在 $\langle\eta\rangle$ 上的正交投影. 令

$$A = I - 2P,$$

则 A 称为关于超平面 $\langle\eta\rangle^{\perp}$ 的镜面反射(n 维线性空间的任一$(n-1)$维子空间称为一个超平面)，简称为镜面反射. 证明：镜面反射是正交变换，并且是第二类的.

3. 设 A 是 n 维欧几里得空间 V 上的一个正交变换，并且 1 是 A 的一个特征值，A 的属于 1 的特征子空间 V_1 的维数是 $n-1$. 证明：A 是镜面反射.

4. 证明：实内积空间 V 到自身的满射 A 是正交变换当且仅当 A 是保持向量长度不变的线性变换.

5. 设 V 是二维欧几里得空间，A 是 V 上的一个正交变换，证明：A 在 V 的任意一个标准正交基下的矩阵 A 是

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \text{ 或 } \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix},$$

其中 θ 是某个实数.

6. 设 α,β 是欧几里得空间 V 中两个不同的单位向量，证明：存在一个镜面反射 A，使得 $A\alpha=\beta$.

7. 证明：正交矩阵的特征多项式的复根为 1 或 -1 或 $\cos\theta\pm i\sin\theta$，$\theta$ 为实数. 且 $0<\theta<2\pi$.

8. 用空间分解的方法证明定理 7.

§5 酉 空 间

在复数域上的线性空间 V 中，如何引进度量概念？关键是要引进内积的概念. 能不能照搬实数域上线性空间的内积的定义？如果照搬，那么对于 $\alpha\neq 0$，有

$$(i\alpha, i\alpha) = i^2(\alpha,\alpha) = -(\alpha,\alpha) < 0,$$
$$(i\alpha, i\alpha) > 0,$$

这是矛盾的. 修改的办法是把对称性用下述性质代替：

$$(\alpha,\beta) = \overline{(\beta,\alpha)}.$$

这一性质称为埃尔米特(Hermite)性，于是

$$(i\alpha, i\alpha) = i(\alpha, i\alpha) = i\overline{(i\alpha,\alpha)} = i\bar{i}(\alpha,\alpha) = (\alpha,\alpha) > 0,$$

这样就没有矛盾了.

定义 1 复数域上线性空间 V 上的一个二元函数记作(α,β)，如果它满足下述 4 条性质：$\forall\, \alpha,\beta,\gamma\in V$，$k\in C$，有

1° $(\alpha,\beta) = \overline{(\beta,\alpha)}$ (埃尔米特性)；

2° $(\alpha+\gamma,\beta) = (\alpha,\beta)+(\gamma,\beta)$ (线性性之一)；

3° $(k\alpha,\beta)=k(\alpha,\beta)$ (线性性之二);

4° (α,α)是非负实数,$(\alpha,\alpha)=0$ 当且仅当 $\alpha=0$ (正定性),

则称这个二元函数(α,β)是 V 上的一个内积.复线性空间 V 上如果指定了一个内积,则称 V 是酉空间.

例 1 C^n 中,对于任意 $X=(x_1,x_2,\cdots,x_n)$, $Y=(y_1,y_2,\cdots,y_n)$,规定

$$(X,Y)\overset{\text{def}}{=}x_1\overline{y_1}+x_2\overline{y_2}+\cdots+x_n\overline{y_n},\tag{1}$$

容易验证,(X,Y)是 C^n 上的一个内积,这个内积称为 C^n 上的标准内积.C^n 装备了这个标准内积,便成为一个酉空间.

例 2 用 $C[a,b]$表示区间$[a,b]$上所有连续复值函数组成的线性空间,规定

$$(f(x),g(x))\overset{\text{def}}{=}\int_a^b f(x)\overline{g(x)}\,dx,\tag{2}$$

容易验证,$(f(x),g(x))$是 $C[a,b]$上的一个内积,此时 $C[a,b]$成为一个酉空间.

例 3 我们用 A^* 表示矩阵 A 中所有元素取共轭复数后再转置(即 $\overline{A}'$).在 $M_n(C)$中,规定

$$(A,B)\overset{\text{def}}{=}\operatorname{tr}(AB^*),\tag{3}$$

容易验证,(A,B)是 $M_n(C)$上的一个内积.此时 $M_n(C)$成为一个酉空间.

与实内积空间类似,酉空间 V 中由于有了内积的概念,从而就有长度、角度、正交、距离等度量概念.

定义 2 非负实数 $\sqrt{(\alpha,\alpha)}$称为向量 α 的长度,记作$|\alpha|$(或者$\|\alpha\|$).

显然$|0|=0$;当 $\alpha\neq 0$ 时,$|\alpha|>0$.容易证明:

$$|k\alpha|=|k||\alpha|,\quad \forall\,\alpha\in V,k\in C.\tag{4}$$

定理 1(Cauchy - Buniakowski 不等式) 在酉空间 V 中,对于任意向量 α, β,有

$$|(\alpha,\beta)|\leq|\alpha||\beta|,\tag{5}$$

等号成立当且仅当 α,β 线性相关.

证明 当 α,β 线性相关时,与实内积空间的情形一样,可证出

$$|(\alpha,\beta)|=|\alpha||\beta|.$$

如果 α,β 线性无关,则对任意复数 t,有 $\alpha+t\beta\neq 0$.从而

$$0<|\alpha+t\beta|^2=|\alpha|^2+\bar{t}(\alpha,\beta)+t(\beta,\alpha)+t\bar{t}|\beta|^2.\tag{6}$$

特别地,取 $t=-\dfrac{(\alpha,\beta)}{|\beta|^2}$,代入(6)式得

$$0<|\alpha|^2-\frac{|(\alpha,\beta)|^2}{|\beta|^2}-\frac{|(\alpha,\beta)|^2}{|\beta|^2}+\frac{|(\alpha,\beta)|^2}{|\beta|^2}$$

$$=|\alpha|^2-\frac{|(\alpha,\beta)|^2}{|\beta|^2},$$

由此得出, $|(\alpha,\beta)| < |\alpha||\beta|$. x

定义 3 酉空间 V 中,两个非零向量 α,β 的夹角 $\langle\alpha,\beta\rangle$ 规定为

$$\langle\alpha,\beta\rangle \stackrel{\text{def}}{} \arccos\frac{|(\alpha,\beta)|}{|\alpha||\beta|}, \tag{7}$$

于是

$$0\leq\langle\alpha,\beta\rangle\leq\frac{\pi}{2}.$$

从(7)式得出, $\langle\alpha,\beta\rangle=\frac{\pi}{2} \quad (\alpha,\beta)=0$.

定义 4 在酉空间 V 中,如果 $(\alpha,\beta)=0$,则称 α 与 β 正交,记作 $\alpha\perp\beta$.

与实内积空间一样,我们可以证明在酉空间中,有三角形不等式和勾股定理.我们可以定义两个向量 α,β 的距离

$$d(\alpha,\beta) \stackrel{\text{def}}{} |\alpha-\beta|.$$

与实内积空间一样,在酉空间 V 中,有正交向量组的概念,并且可以证明:正交向量组一定线性无关.从而有正交基、标准正交基的概念,利用施密特正交化和单位化,可把 V 的一个基变成与它等价的标准正交基.

n 维酉空间 V 中,向量组 $\eta_1,\eta_2,\cdots,\eta_n$ 是 V 的一个标准正交基当且仅当

$$(\eta_i,\eta_j)=\delta_{ij},\quad i,j=1,2,\cdots,n. \tag{8}$$

利用标准正交基 $\eta_1,\eta_2,\cdots,\eta_n$,容易计算向量的内积.设 α,β 在 $\eta_1,\eta_2,\cdots,\eta_n$ 下的坐标分别是 $X=(x_1,x_2,\cdots,x_n)'$, $Y=(y_1,y_2,\cdots,y_n)'$,则

$$\begin{aligned}(\alpha,\beta)&=\left(\sum_{i=1}^{n}x_i\eta_i,\ \sum_{j=1}^{n}y_j\eta_j\right)\\&=\sum_{i=1}^{n}\sum_{j=1}^{n}x_i\bar{y}_j(\eta_i,\eta_j)\\&=\sum_{i=1}^{n}x_i\bar{y}_i=Y^*X.\end{aligned} \tag{9}$$

利用标准正交基,向量的坐标的分量可以用内积表达.设 α 在标准正交基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的坐标是 $(x_1,x_2,\cdots,x_n)'$,则 $\alpha=\sum_{i=1}^{n}x_i\eta_i$.两边用 η_j 作内积,得

$$(\alpha,\eta_j)=\sum_{i=1}^{n}x_i(\eta_i,\eta_j)=x_j,$$

因此

$$\alpha=\sum_{i=1}^{n}(\alpha,\eta_i)\eta_i. \tag{10}$$

(10)式称为 α 的傅里叶(Fourier)展开,其中每个系数 (α,η_i) 称为 α 的傅里叶

(Fourier)系数.

n 维酉空间 V 中,设 $\eta_1,\eta_2,\cdots,\eta_n$ 是 V 的一个标准正交基,向量组 β_1,$\beta_2,\cdots,\beta_n$ 满足

$$(\beta_1,\beta_2,\cdots,\beta_n)=(\eta_1,\eta_2,\cdots,\eta_n)P, \tag{11}$$

则 β_i 在标准正交基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的坐标是 P 的第 i 列 X_i, $i=1,2,\cdots,n$,于是

$$\beta_1,\beta_2,\cdots,\beta_n \text{ 是 } V \text{ 的一个标准正交基}$$

$$(\beta_i,\beta_j)=\delta_{ij},\ i,j=1,2,\cdots,n$$

$$X_j^* X_i=\delta_{ij},\ i,j=1,2,\cdots,n$$

$$\begin{matrix} X_1^* X_1 & X_1^* X_2 & \cdots & X_1^* X_n \\ X_2^* X_1 & X_2^* X_2 & \cdots & X_2^* X_n \\ \cdots & \cdots & & \cdots \\ X_n^* X_1 & X_n^* X_2 & \cdots & X_n^* X_n \end{matrix} = \begin{matrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & & \cdots \\ 0 & 0 & \cdots & 1 \end{matrix}$$

$$\begin{matrix} X_1^* \\ X_2^* \\ \cdots \\ X_n^* \end{matrix} (X_1,X_2,\cdots,X_n)=I$$

$$P^* P=I. \tag{12}$$

定义 5　复数域上 n 级矩阵 P 如果满足

$$P^* P=I,$$

则称 P 是酉矩阵.

从定义 5 得出

$$n \text{ 级复矩阵 } P \text{ 是酉矩阵} \qquad P^* P=I$$

$$P \text{ 可逆,且 } P^{-1}=P^*$$

$$PP^*=I.$$

上面的讨论表明,n 维酉空间 V 中,标准正交基到标准正交基的过渡矩阵是酉矩阵.反之,如果向量组 $\beta_1,\beta_2,\cdots,\beta_n$ 满足(11)式.且 P 是酉矩阵,则 β_1,$\beta_2,\cdots,\beta_n$ 是 V 的一个标准正交基.

与实内积空间的情形一样,酉空间有同构的概念,并且同样可以证明:两个有限维酉空间同构的充分必要条件是它们的维数相同.

与实内积空间的情形一样,酉空间中有正交补的概念,并且同样可以证明:n 维酉空间 V 等于它的任一子空间 U 与 $U^{\perp}$ 的直和:$V=U \quad U^{\perp}$.从而在 n 维酉空间 V 中,有 V 在 U 上的正交投影,向量 α 在 U 上的正交投影等概念,并且同样可以证明:$\alpha_1 \in U$ 是 α 在 U 上的正交投影当且仅当

$$d(\alpha,\alpha_1)\leq d(\alpha,\gamma),\ \forall\gamma\in U. \tag{13}$$

类似于实内积空间上的正交变换,在酉空间中是酉变换.

定义 6 酉空间 V 到自身的满射 $\mathcal{A}$ 如果保持内积不变,即

$$(\mathcal{A}\alpha,\mathcal{A}\beta)=(\alpha,\beta),\ \forall\alpha,\beta\in V,$$

则称 $\mathcal{A}$ 是 V 上的一个酉变换.

与正交变换的情形一样,可以证明:

命题 2 酉空间 V 上的酉变换一定是线性变换,并且是单射,从而是可逆的.

于是,酉空间 V 上的变换 $\mathcal{A}$ 是酉变换当且仅当 $\mathcal{A}$ 是 V 到自身的一个同构映射,从而酉变换的逆变换还是酉变换;酉变换的乘积还是酉变换.

与正交变换的情形一样,可以证明

命题 3 n 维酉空间 V 上的线性变换 $\mathcal{A}$ 是酉变换

$\mathcal{A}$ 把 V 的标准正交基映成标准正交基

$\mathcal{A}$ 在 V 的标准正交基下的矩阵是酉矩阵.

习 题 10.5

1. 在酉空间 C^3 (指定标准内积)中,设

$$\alpha=(1,-1,1),\ \beta=(1,0,i),$$

求$|\alpha|,|\beta|$,α与β的夹角$\langle\alpha,\beta\rangle$.

2. 在酉空间 C^2 (指定标准内积)中,设

$$\alpha_1=(1,-1),\ \alpha_2=(1,i),$$

求与 α_1,α_2 等价的一个标准正交基 η_1,η_2.

3. 在酉空间 C^3 (指定标准内积)中,设

$$\alpha_1=(1,-1,1),\ \alpha_2=(1,0,i),$$

求与 α_1,α_2 等价的一个正交向量组 β_1,β_2.

4. 写出 1 级酉矩阵的形式.

5. 证明:酉矩阵的行列式的模为 1.

6. 证明:酉变换的特征值的模为 1.

7. n 级复矩阵 A 如果满足

$$A^*=A,$$

则称 A 是埃尔米特(Hermite)矩阵(或者自伴矩阵),写出 2 级 Hermite 矩阵的形式.

*8. 酉空间 V 上的线性变换 $\mathcal{A}$ 如果满足

$$(\mathcal{A}\alpha,\beta)=(\alpha,\mathcal{A}\beta),\ \forall\alpha,\beta\in V,$$

则称 $\mathcal{A}$ 是埃尔米特(Hermite)变换(或者自伴变换).证明:n 维酉空间 V 上的线性变换 $\mathcal{A}$ 是

Hermite 变换当且仅当 $\mathcal{A}$ 在 V 的标准正交基下的矩阵是 Hermite 矩阵.

*9. 证明:酉空间 V 上的 Hermite 变换 $\mathcal{A}$ 的特征值一定是实数.

*§6　正交空间与辛空间

我们已经分别在实数域上和复数域上的线性空间中,通过引进内积的概念,使得这些空间中有长度、角度、正交、距离等度量概念.对于任意域 F 上的线性空间 V 中,能不能也引进度量概念?即便是对于实数域上线性空间 V,在有的实际问题里,也不用正定对称双线性函数引进度量概念.例如,作为爱因斯坦(Einstein)相对论基础的"时-空"空间,即闵可夫斯基(Minkowski)空间 V,它是实数域上的4维线性空间,并且指定了一个非退化的对称双线性函数 f 作为度量,也称为内积.这样做的目的是使得洛伦兹变换保持 V 上的内积不变,从而保持时-空间隔的平方(即,$f(\alpha-\beta,\alpha-\beta)$,其中 $\alpha,\beta\in V$)不变.

定义1　域 F 上线性空间 V 如果指定了一个对称双线性函数 f,则称 V 是一个正交空间,称 f 是 V 上的一个内积(或度量).用(V,f)表示指定的内积为 f 的正交空间,如果 f 是非退化的,则(V,f)称为正则的;否则称为非正则的.

例如,在闵可夫斯基空间中,内积 f 是

$$f(\alpha,\beta)=x_1y_1+x_2y_2+x_3y_3-c^2t_1t_2, \tag{1}$$

其中 c 是光速,$\alpha=(t_1,x_1,x_2,x_3)$,$\beta=(t_2,y_1,y_2,y_3)$.

在正交空间中,由于内积 f 不要求有正定性,因此就无法引进长度、角度、距离等度量概念,但是仍有正交这一概念.

定义2　在正交空间(V,f)中,如果 $f(\alpha,\beta)=0$,则称 α 与 β 正交,记作 $\alpha\perp\beta$.

由于 f 是对称的,于是从 $f(\alpha,\beta)=0$ 可推出 $f(\beta,\alpha)=0$,从而如果 α 与 β 正交,则 β 与 α 也正交.

在正交空间中,一个非零向量有可能与自身正交.例如,在 $\mathbf{R}^4$ 中,对于 $\alpha=(x_1,x_2,x_3,x_4)$,$\beta=(y_1,y_2,y_3,y_4)$,令

$$f(\alpha,\beta)\overset{\text{def}}{=}x_1y_1-x_2y_2-x_3y_3-x_4y_4, \tag{2}$$

容易验证,f 是 $\mathbf{R}^4$ 上一个非退化的对称双线性函数.在正交空间$(\mathbf{R}^4,f)$中,设 $\alpha_1=(1,1,0,0)$,则

$$f(\alpha_1,\alpha_1)=1^2-1^2=0,$$

从而 α_1 与自身正交.这样的非零向量称为迷向向量.

定义3　设 S 是正交空间(V,f)的一个子集,下述集合

$$\{\alpha\in V \mid f(\alpha,\beta)=0,\ \forall\ \beta\in S\},$$

称为 S 的正交补,记作 $S^{\perp}$.

容易看出,$S^{\perp}$ 是 V 的线性子空间.

定理 1 设(V, f)是有限维正则的正交空间,W 是 V 的一个子空间,则

(i) $\dim W+\dim W^{\perp}=\dim V$; (3)

(ii) $(W^{\perp})^{\perp}=W$. (4)

证明 这实际上就是习题 10.1 的第 8 题,已经给出了证明的提示,这里不再重复写出. x

定义 4 有限维正交空间(V, f)的一个基 $\alpha_1,\cdots,\alpha_n$ 称为正交基,如果 $\alpha_1,\cdots,\alpha_n$ 两两正交,即 $f(\alpha_i,\alpha_j)=0$,对一切 $i\neq j$.

注意正交基的定义中,首先要求是基,然后要求基向量两两正交.这是因为在正交空间中两两正交的向量可能是线性相关的.例如,在例 1 的(R^4, f)中,设 $\alpha=(1,1,0,0)$,则 α 与 2α 是正交的,然而它们是线性相关的.

定理 2 特征不为 2 的域 F 上的有限维的正交空间(V, f)一定存在正交基.

证明 据§1 的定理 4,存在 V 的一个基 $\alpha_1,\cdots,\alpha_n$,使得 f 在此基下的度量矩阵 A 是对角矩阵 $\mathrm{diag}\{d_1,\cdots,d_n\}$.这表明 $\alpha_1,\cdots,\alpha_n$ 两两正交,因此它是 V 的一个正交基. x

注意当(V, f)是正则的时候,上述证明中的 A 是满秩矩阵,从而 $d_i\neq 0$,$i=1,\cdots,n$.于是 $f(\alpha_i,\alpha_i)\neq 0$,这表明 α_i 是非迷向的,$i=1,\cdots,n$.这也就是说,当(V, f)正则时,存在由非迷向向量组成的正交基.

定义 5 有限维正交空间(V, f)的一个正交基 $\varepsilon_1,\cdots,\varepsilon_n$,称为标准正交基,如果

$$f(\varepsilon_i,\varepsilon_i)=0 \text{ 或 } \pm 1,\ i=1,\cdots,n.$$

当 $F=R$ 或 C 时,有限维正交空间(V, f)存在标准正交基,并且当 $F=C$ 时,总可以使标准正交基中的每个向量 ε_i,满足

$$f(\varepsilon_i,\varepsilon_i)=0 \text{ 或 } 1.$$

对于有限维的正则的正交空间(V, f),它的每个向量 β 在 V 的正交基 $\alpha_1,\cdots,\alpha_n$ 下的坐标特别简单好记,即我们有

$$\beta=\sum_{i=1}^{n}\frac{f(\beta,\alpha_i)}{f(\alpha_i,\alpha_i)}\alpha_i, \tag{5}$$

公式(5)的证明留作习题.

定理 3 设 $\mathrm{char}\,F\neq 2$,设(V, f)是正交空间,如果 W 是 V 的有限维正则子空间,则

$$V = W \oplus W^{\perp}. \tag{6}$$

证明　任取 $\alpha \in W \cap W^{\perp}$.因为 W 是正则的,所以 $\operatorname{rad} W = 0$,从而 $\alpha = 0$.因此 $W \cap W^{\perp} = 0$.

再证 $V = W + W^{\perp}$.在 W 中取一个由非迷向向量组成的正交基 $\alpha_1, \cdots, \alpha_m$,对于 V 中任一向量 β,令

$$\beta_1 = \sum_{i=1}^{m} \frac{f(\beta, \alpha_i)}{f(\alpha_i, \alpha_i)} \alpha_i \in W, \quad \beta_2 = \beta - \beta_1.$$

因为
$$\begin{aligned} f(\beta_2, \alpha_j) &= f(\beta, \alpha_j) - \sum_{i=1}^{m} \frac{f(\beta, \alpha_i)}{f(\alpha_i, \alpha_i)} f(\alpha_i, \alpha_j) \\ &= f(\beta, \alpha_j) - f(\beta, \alpha_j) = 0, \quad j = 1, \cdots, m, \end{aligned}$$

所以 $\beta_2 \in W^{\perp}$,于是

$$\beta = \beta_1 + \beta_2 \in W + W^{\perp},$$

所以 $V = W \oplus W^{\perp}$.　□

定义 6　特征不为 2 的域 F 上线性空间 V 如果指定了一个斜对称双线性函数 f,则称 V 是一个辛空间,用 (V, f) 表示,称 f 是 V 上的一个内积(或辛内积);如果 f 是非退化的,则称 (V, f) 是正则的;否则称为非正则的.

例如,R^2 中,对于 $\alpha = (x_1, x_2)$,$\beta = (y_1, y_2)$,令

$$f(\alpha, \beta) \overset{\text{def}}{=} x_1 y_2 - x_2 y_1, \tag{7}$$

容易验证,f 是 R^2 上一个非退化的斜对称双线性函数.于是 (R^2, f) 成为一个辛空间.

与正交空间一样,辛空间中有正交的概念,但没有长度、角度、距离等概念,辛空间中也有迷向向量.

从§1 的公式 20 知道,有限维的正则的辛空间一定是偶数维的.

从§1 的定理 5 知道,n 维辛空间 (V, f) 中,存在 V 的一个基 $\varepsilon_1, \varepsilon_{-1}, \cdots, \varepsilon_r, \varepsilon_{-r}, \eta_1, \cdots, \eta_s$,使得 f 在这个基下的度量矩阵 A 为

$$\operatorname{diag}\left\{ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \cdots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, 0, \cdots, 0 \right\},$$

即
$$\begin{aligned} & f(\varepsilon_i, \varepsilon_{-i}) = 1, \quad i = 1, \cdots, r, \\ & f(\varepsilon_i, \varepsilon_j) = 0, \quad i + j \neq 0, \\ & f(\varepsilon_i, \eta_k) = 0, \quad i = \pm 1, \cdots, \pm r, \\ & f(\eta_j, \eta_k) = 0, \quad j, k = 1, \cdots, s, \end{aligned} \tag{8}$$

这个基称为 (V, f) 的辛基.

我们把基 $\varepsilon_1, \cdots, \varepsilon_r, \varepsilon_{-1}, \cdots, \varepsilon_{-r}, \eta_1, \cdots, \eta_s$ 也称为辛基.容易看出,f 在此基下的度量矩阵为

$$\begin{matrix} 0 & I_r & 0 \\ -I_r & 0 & 0 \\ 0 & 0 & 0 \end{matrix}.$$

定理 4 设(V, f)是 n 维辛空间,如果 W 是 V 的正则子空间,则

$$V = W \oplus W^{\perp}. \tag{9}$$

证明 任取 $\alpha \in W \cap W^{\perp}$,因为 W 是正则的,所以 $\mathrm{rad}\, W = 0$,从而 $\alpha = 0$.因此 $W \cap W^{\perp} = 0$.

在 W 中取一个辛基 $\varepsilon_1, \varepsilon_{-1}, \cdots, \varepsilon_m, \varepsilon_{-m}$.把它扩充成 V 的一个基:$\varepsilon_1, \varepsilon_{-1}, \cdots, \varepsilon_m, \varepsilon_{-m}, \beta_1, \cdots, \beta_t$,令

$$\beta_j' = \beta_j - \sum_{i=1}^{m} f(\beta_j, \varepsilon_{-i})\varepsilon_i + \sum_{i=1}^{m} f(\beta_j, \varepsilon_i)\varepsilon_{-i}, \quad j = 1, \cdots, t,$$

则

$$\begin{aligned} f(\beta_j', \varepsilon_k) &= f(\beta_j, \varepsilon_k) - \sum_{i=1}^{m} f(\beta_j, \varepsilon_{-i}) f(\varepsilon_i, \varepsilon_k) \\ &\quad + \sum_{i=1}^{m} f(\beta_j, \varepsilon_i) f(\varepsilon_{-i}, \varepsilon_k) \\ &= f(\beta_j, \varepsilon_k) + f(\beta_j, \varepsilon_k) f(\varepsilon_{-k}, \varepsilon_k) = 0, \ k = 1, \cdots, m, \end{aligned}$$

$$f(\beta_j', \varepsilon_{-k}) = f(\beta_j, \varepsilon_{-k}) - f(\beta_j, \varepsilon_{-k}) f(\varepsilon_k, \varepsilon_{-k}) = 0, \ k = 1, \cdots, m.$$

显然 $\varepsilon_1, \varepsilon_{-1}, \cdots, \varepsilon_m, \varepsilon_{-m}, \beta_1', \cdots, \beta_t'$仍是 V 的一个基.

令

$$U = \langle \beta_1', \cdots, \beta_t' \rangle.$$

则 $V = W \oplus U$,由上面推导出的结果知道 $\beta_j' \in W^{\perp}$, $j = 1, \cdots, t$.所以 $U \subseteq W^{\perp}$.另一方面,任取 $\alpha \in W^{\perp}$,设

$$\alpha = \alpha_1 + \alpha_2, \quad \alpha_1 \in W, \quad \alpha_2 \in U,$$

任取 $\beta \in W$,有

$$0 = f(\alpha, \beta) = f(\alpha_1, \beta) + f(\alpha_2, \beta) = f(\alpha_1, \beta),$$

于是 $\alpha_1 \in W^{\perp}$.又 $\alpha_1 \in W$,所以 $\alpha_1 = 0$.从而 $\alpha \in U$.因此 $U = W^{\perp}$.于是有 $V = W \oplus W^{\perp}$. ∎

*习 题 10.6

1. R^2 中,对于 $\alpha' = (x_1, x_2)$, $\beta' = (y_1, y_2)$,定义

$$f(\alpha, \beta) = x_1 y_1 - x_2 y_2.$$

(1) 证明:(R^2, f)是一个正则的正交空间;

(2) 设 $\varepsilon_1' = (1,0)$, $\varepsilon_2' = (0,1)$,证明:$\varepsilon_1, \varepsilon_2$ 是(R^2, f)的一个标准正交基;

(3) 求 f 在基 $\varepsilon_1, \varepsilon_2$ 下的度量矩阵 A;

(4) 设

$$T = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

令 $\mathcal{T}$ 是 R^2 上的一个线性变换,它在基 $\varepsilon_1, \varepsilon_2$ 下的矩阵为 T.证明:$\mathcal{T}$ 是正交空间(R^2, f)上的一个正交变换(域 F 上有限维正则的正交空间 V 上的一个线性变换 $\mathcal{T}$ 如果保持内积不变,则称 $\mathcal{T}$ 是 V 上的一个正交变换),并且它的特征值是 2 ± 1;

(5) 求 $\mathcal{T}$ 的全部特征向量,并且说明 $\mathcal{T}$ 的特征向量都是迷向的.

2. 证明本节公式(5).

3. 设(V, f)是有限维正则辛空间,V 上的一个线性变换 $\mathcal{B}$ 如果保持辛内积不变,则称 $\mathcal{B}$ 为 V 上的辛变换.证明:V 上的线性变换 $\mathcal{B}$ 是辛变换当且仅当 $\mathcal{B}$ 在 V 的辛基 $\varepsilon_1, \cdots, \varepsilon_r, \varepsilon_{-1}, \cdots, \varepsilon_{-r}$ 下的矩阵 B 满足 $B'AB = A$,其中

$$A = \begin{pmatrix} 0 & \mathbf{I}_r \\ -\mathbf{I}_r & 0 \end{pmatrix}.$$

满足 $B'AB = A$ 的矩阵 B 称为辛矩阵.

习题答案与提示

第7章　多项式环

习　题　7.1

1. $f(x)$与$g(x)$的次数相等.

2. $f(x)$与$g(x)$的次数都是0.

3. $f(x)+g(x)$的次数不一定是3,可能小于3.例如,设$f(x)=x^3+1$,$g(x)=-x^3+x^2-1$,则$f(x)+g(x)$的次数是2.

4. 提示:$f(x)$是可逆元当且仅当存在$g(x)\in K[x]$,使得,$f(x)g(x)=1$.然后用多项式乘积的次数公式.

*5. 提示:用反证法.

6. 提示:设

$$D=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix},$$

则

$$A=I+bD+b^2D^2+\cdots+b^{n-1}D^{n-1}.$$

然后利用一元多项式环的通用性质.

7. 提示:设$A=(a_{ij})$,由行列式的定义得

$$|\lambda I-A|=\sum_{j_1j_2\cdots j_n}(-1)^{\tau(j_1j_2\cdots j_n)}(\lambda\delta_{1j_1}-a_{1j_1})\cdots(\lambda\delta_{nj_n}-a_{nj_n}).$$

然后用一元多项式环的通用性质(λ用$k^{-1}\lambda$代入).

*8. 提示:$|\lambda I+A||\lambda I-A|=|\lambda^2I-A^2|$.设$A^2=(b_{ij})$.

则

$$|\lambda^2I-A^2|=\sum_{j_1j_2\cdots j_n}(-1)^{\tau(j_1j_2\cdots j_n)}(\lambda^2\delta_{1j_1}-b_{1j_1})\cdots(\lambda^2\delta_{nj_n}-b_{nj_n}).$$

然后用第7题的已知条件和结论(令$k=-1$),以及对于$K[\lambda^2]$运用一元多项式环的通用性质(λ^2用λ代入).

习　题　7.2

1. 用整除的定义可推导出结论.

2. 用整除的定义可推导出结论.

3. (1) 商式是 x^2+2x-4,余式是 $-20x+19$;

(2) 商式是$\frac{1}{3}x^2+\frac{4}{9}x-\frac{2}{27}$,余式是 $-\frac{80}{27}x+\frac{85}{27}$.

4. $g(x)\mid f(x)$当且仅当 $a_1=3$ 且 $a_0=-1$.

5. (1) 商式是 $3x^3+12x^2+43x+174$,余式是 695;

(2) 商式是 $5x^2-10x+17$,余式是 -30.

*6. 用整除的定义可推出(1)～(4)的结论.

习 题 7.3

1. (1) $(f(x),g(x))=x+3$;

(2) $(f(x),g(x))=x-1$.

2. 去证 $f(x)$与 $g(x)$的任一公因式 $c(x)$能整除 $d(x)$.

3. 利用定理 3 和第 2 题的结论.

4. 设 $f(x)=f_1(x)(f(x),g(x))$, $g(x)=g_1(x)(f(x),g(x))$.利用定理 3 和定理 4 可证得,$(f_1(x),g_1(x))=1$.

5. 利用定理 4.

6. 利用定理 4 可证得,$(f,f+g)=1$.同理有,$(g,f+g)=1$.然后利用性质 3 可得,$(fg,f+g)=1$.

7. 去证 $af+bg$ 与 $cf+dg$ 的所有公因式组成的集合等于 f 与 g 的所有公因式组成的集合,然后用命题 1.

8. 用定理 4 以及一元多项式环的通用性质.

9. 必要性.设$(f(x),g(x))=d(x)$,其中 $\deg d(x)>1$.设 $f(x)=f_1(x)d(x)$, $g(x)=g_1(x)d(x)$,可证得结论.

充分性,用反证法.

10. (1) 取 $f(x)$与 $g(x)$的一个最大公因式 $d(x)$,设 $f(x)=f_1(x)d(x)$, $g(x)=g_1(x)d(x)$.令 $m(x)=f_1(x)g_1(x)d(x)$,去证 $m(x)$是 $f(x)$与 $g(x)$的一个最小公倍式.

(2) 从(1)的证明过程看出, $m(x)=\frac{f(x)g(x)}{d(x)}$,由此即可得出结论.

*11. 用定理 3 以及一元多项式环的通用性质可证得:如果 α属于 $f(A)X=0$ 的解空间 W_1 与 $g(A)X=0$ 的解空间 W_2 的交,则 α是 $d(A)X=0$ 的解.设 $f(x)=f_1(x)d(x)$, $g(x)=g_1(x)d(x)$,用一元多项式环的通用性质可证得:如果 γ是 $d(A)X=0$ 的解,则 γ也是 $f(A)X=0$ 的解,并且 γ是 $g(A)X=0$ 的解.

*12. 利用定理 4 以及一元多项式环的通用性质可证出:$f(A)X=0$ 的任一个解可以表示成 $f_1(A)X=0$ 的一个解与 $f_2(A)X=0$ 的一个解的和.唯一性可利用第 11 题的结论证得.

习 题 7.4

1. 可用反证法.

2. (1) 在复数域上:

$$x^4+1= \quad x+\quad \frac{2}{2}-\frac{2}{2}i \qquad x-\quad \frac{2}{2}-\frac{2}{2}i \qquad x+\quad \frac{2}{2}+\frac{2}{2}i \qquad x-\quad \frac{2}{2}+\frac{2}{2}i \quad ;$$

在实数域上:

$$x^4+1=(x^2+\ 2x+1)(x^2-\ 2x+1);$$

在有理数域上: x^4+1 不可约.

(2) 在复数域上:

$$x^4+4=[x+(1-i)][x-(1-i)][x+(1+i)][x-(1+i)];$$

在实数域上:

$$x^4+4=(x^2+2x+2)(x^2-2x+2);$$

在有理数域上:

$$x^4+4=(x^2+2x+2)(x^2-2x+2).$$

3. 充分性是显然的.必要性利用 $f(x)$和 $g(x)$的标准分解式得出 $f^2(x)$和 $g^2(x)$的标准分解式.由 $g^2(x)\mid f^2(x)$可看出 $g^2(x)$的每一个不可约因式与 $f^2(x)$的不可约因式的关系,从而可推出 $g(x)\mid f(x)$.

4. 用反证法,且利用本节性质 3.

*5. 必要性.利用本节性质 1.充分性.用反证法,假设 $f(x)$不是一个不可约多项式的方幂,去证存在一个 $g(x)\in K[x]$使得,$(f(x),g(x))\neq 1$ 且对一切正整数 m 都有 $f(x)8\ g^m(x)$.为此设

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_s^{r_s}(x).$$

由所设,$s\geq 2$.取 $g(x)=p_2(x)$.

*6. 必要性.利用第 5 题的结论.充分性.用反证法.假设 $f(x)$不是一个不可约多项式的方幂,则 $f(x)$的标准分解式为

$$f(x)=cp_1^{r_1}(x)p_2^{r_2}(x)\cdots p_s^{r_s}(x),$$

其中 $s\geq 2$,取 $g(x)=cp_1^{r_1}(x)$, $h(x)=p_2^{r_2}(x)\cdots p_s^{r_s}(x)$.

7. 显然 g_1 与 g_2 的每一个公因式都是 fg_1 与 g_2 的公因式.反之,设 $c(x)$是 fg_1 与 g_2 的公因式,利用§3 的性质 3 和定理 4,去证 $c(x)$也是 g_1 与 g_2 的公因式,然后用§3 的命题 1 便得出结论.

习　题　7.5

1. (1) 有重因式 $x-2$;

(2) 没有重因式.

2. $f(x)=x^3+2ax+b$ 有重因式的充分必要条件是

$$32a^3+27b^2=0.$$

3. 例如 $f(x)=x^k+1(k\geq 2)$,则 $f'(x)=kx^{k-1}$.

4. 设 $p(x)$是 $f(x)$的 m 重因式,则 $p(x)$是 $f'(x)$的 $m-1$ 重因式.由已知条件得,$m-1=k-1$,从而 $m=k$.

5. 必要性.由定理 1 得到.充分性.利用第 4 题的结论.

6. $g(x)=x^2-1$, $f(x)=(x+1)(x-1)^4$.

*7. 充分性显然.必要性.利用去掉 $f(x)$的不可约因式的重数的方法可得,

$$f(x)=g(x)(f(x),f'(x)),$$

其中 $g(x)$与 $f(x)$含有完全相同的不可约因式,但 $g(x)$没有重因式.因为 $f'(x)\mid f(x)$,所以$(f(x),f'(x))=cf'(x)$.从而 $f(x)=cg(x)f'(x)$.比较此式两边的多项式的次数得 $\deg g(x)=1$,于是 $g(x)=a(x-b)$, $a,b\in K$.从 $g(x)$与 $f(x)$的不可约因式的关系立即得出所要证的结论.

习 题 7.6

1. -2 是 $f(x)$的 3 重根.

2. $f(x)$有重根当且仅当 $a=4$ 或 $a=-\frac{17}{27}$.

当 $a=4$ 时,2 是 $f(x)$的二重根;当 $a=-\frac{17}{27}$时,$\frac{1}{3}$是 $f(x)$的二重根.

3. $f(x)$与 $g(x)$只有一个公共根:2.

4. 3 是 $f(x)$的二重根当且仅当 $a=4$ 且 $b=3$.

5. 利用一次因式与根的关系,以及一元多项式环的通用性质.

6. 由整除的定义,以及一元多项式环的通用性质 x 分别用 ω,ω^2 代入,其中 $\omega=\frac{-1+\sqrt{3}i}{2}$ 可得到关于 $f_1(1)$和 $f_2(1)$的两个方程.

7. 由整除的定义,以及一元多项式环的通用性质可证得:若 c是 $f(x)$的一个复根,则 c^m 也是 $f(x)$的一个复根.

8. 令 $\xi=e^{i\frac{2\pi}{n}}$,则 $1,\xi,\xi^2,\cdots,\xi^{n-1}$都是 n次单位根,且它们两两不同,因此

$$x^n-1=(x-1)(x-\xi)(x-\xi^2)\cdots(x-\xi^{n-1}).$$

*9. (1) 取 $m(x)$是 J_a 中次数最低的首项系数为 1 的多项式,对于 $f(x)\in J_a$,用 $m(x)$去除 $f(x)$,证明所得的余式必为零多项式.

(2) 用反证法,利用§4 的不可约多项式的性质 3,以及一元多项式环的通用性质.

10. 数域 K 上不可约多项式 $p(x)$在 $K[x]$中没有重因式,从而在 $C[x]$中,$p(x)$也没有重因式.因此 $p(x)$在复数域内没有重根.

11. 据本套教材上册第 176 页的命题 2,A 的特征多项式为

$$f(\lambda)=|\lambda I-A|=\lambda^n-\mathrm{tr}(A)+\cdots+(-1)^n|A|.$$

然后用 Vieta 公式.

12. 去证 $p(x)$与 $g(x)$在 $K[x]$中不互素.注意利用互素性不随数域的扩大而改变.

习 题 7.7

1. 因为实系数多项式的虚根共轭成对出现.

2. 设 $\xi=e^{i\frac{2\pi}{n}}$,当 $1\leq j<n$ 时,ξ^j 与ξ^{n-j}是一对共轭复数.因此当 $n=2m+1$ 时,

$$x^n-1=(x-1)\left(x^2-2\cos\frac{2\pi}{n}x+1\right)\left(x^2-2\cos\frac{4\pi}{n}x+1\right)\cdots\left(x^2-2\cos\frac{2m\pi}{n}x+1\right);$$

当 $n=2m$ 时,

$$x^n-1=(x-1)(x+1)\left(x^2-2\cos\frac{2\pi}{n}x+1\right)\cdots\left(x^2-2\cos\frac{2(m-1)\pi}{n}x+1\right).$$

3. (1) 把 A 看成复矩阵,则 $f(\lambda)$的任一复根 λ_0 是 A 的特征值,从而存在 n维复向量 ξ 使得 $A\xi=\lambda_0\xi$.注意一个复数 z 为纯虚数或零当且仅当 $\bar{z}=-z$.因此只要去证明 $\bar{\lambda}_0+\lambda_0=0$.为此可以仿照本套教材上册第 183 页的定理 1 的证法.

(2) 由第(1)小题和实系数多项式唯一因式分解定理立即得出结论.

*4. 用反证法.假如 A 和 B 作为复矩阵相似,则存在复数域上的 n级可逆矩阵 U 使得,$U^{-1}AU=B$.设 $U=P+iQ$,其中 P, Q 都是实矩阵.从 P, Q 出发,想构造一个可逆实矩阵 S.为此任给实数 t,考虑行列式$|P+tQ|$,记作 $g(t)$,它是 t的至多n次的多项式.从这里可以知道可逆实矩阵 S 应怎么构造.

习　题　7.8

1. (1) $\frac{1}{2}$;　　　　(2) $-\frac{1}{2}$,3.

2. (1) 不可约;(2) 不可约;

(3) x用 $x-1$ 代入,然后用 Eisenstein 判别法可得,不可约;

(4) x用 $x+1$ 代入,不可约;

(5) x用 $x-1$ 代入,不可约;

(6) 没有有理根,且是 3 次多项式,从而不可约;

(7) 有有理根 $-\frac{1}{2}$,从而可约;

(8) 没有有理根;并且用反证法可证得,没有二次因式,因此不可约.

3. 令 $c=\sqrt[n]{p_1p_2\cdots p_n}$,其中 $p_1,p_2,\cdots,p_n$ 是两两不同的素数.则 $c^n=p_1p_2\cdots p_n$.假如 c 是有理数,则多项式 $x^n-p_1p_2\cdots p_n$ 有一个有理根 c,从而有一个一次因式.由于 $n>1$,于是 $x^n-p_1p_2\cdots p_n$ 在 Q 上可约.但是用 Eisenstein 判别法可得,$x^n-p_1p_2\cdots p_n$ 在 Q 上不可约,矛盾.

4. $f(1)=a_n+\cdots+a_1+a_0$ 为奇数,因此 1 不是 $f(x)$的根.假如 -1 是 $f(x)$的根,则 $x+1$是 $f(x)$在 $Q[x]$中的因式.设 $f(x)=rf_1(x)$,其中 $f_1(x)$是本原多项式,$r\in Z^*$,则 $x+1$也是 $f_1(x)$在 $Q[x]$中的因式.据定理 3 和高斯引理得,$f_1(x)=(x+1)h(x)$,其中 $h(x)$是本原多项式.于是 $f(x)=r(x+1)h(x)$.从而 $f(1)=2rh(1)$,这与 $f(1)$是奇数矛盾,因此 -1不是 $f(x)$的根.

证明 -1 不是 $f(x)$ 的根的另一种方法:假如 -1 是 $f(x)$ 的根,则 $a_n(-1)^n+a_{n-1}(-1)^{n-1}+\cdots+a_1(-1)+a_0=0$.从而

$$a_n+a_{n-1}+\cdots+a_1+a_0=\begin{cases}2(a_1+a_3+\cdots+a_{2m-1}), & n=2m,\\ 2(a_1+a_3+\cdots+a_{2m+1}), & n=2m+1.\end{cases}$$

这与 $a_n+a_{n-1}+\cdots+a_1+a_0$ 是奇数矛盾.

5. 假如 $f(x)$有有理根 c.由于 $f(x)$的首项系数为1,因此 $f(x)$是本原多项式,并且据定理4得,c必为整数.于是 $x-c$是 $f(x)$的因式.据定理3和高斯引理得,$f(x)=(x-c)h(x)$,其中 $h(x)$是本原多项式,从而

$$f(0)=-ch(0),\quad f(1)=(1-c)h(1).$$

由于 $-c$与$(-c+1)$必有一个是偶数,因此 $f(0)$与 $f(1)$必有一个是偶数,矛盾.

6. 从$(a+b)c$是奇数可推出,$f(0)$与 $f(1)$都是奇数,然后用第5题的结论.

7. 若 $n=1$,显然 $f(x)$不可约.下面设 $n>1$,假如 $f(x)$在 Q 上可约,则据推论2得,存在 $f_1(x),f_2(x)\in Z[x]$使得

$$f(x)=f_1(x)f_2(x),\quad \deg f_i(x)<\deg f(x),i=1,2.$$

于是

$$f_1(a_i)f_2(a_i)=f(a_i)=1,\quad i=1,2,\cdots,n,$$

从而 $f_1(a_i)$与 $f_2(a_i)$同为1或同为-1,由此得出

$$f_1(a_i)-f_2(a_i)=0,\quad i=1,2,\cdots,n.$$

由于

$$\deg(f_1(x)-f_2(x))\leq\max\{\deg f_1(x),\deg f_2(x)\}<\deg f(x)=n,$$

因此 $f_1(x)-f_2(x)=0$,从而 $f_1(x)=f_2(x)$.于是 $f(x)=f_1^2(x)$.这表明$\deg f(x)$是偶数.因此,如果 n是奇数,则 $f(x)$在 Q 上不可约.

$n=2$时,设 $f(x)=(x-1)(x+1)+1=x^2$ 在 Q 上可约.

$n=4$时,设 $f(x)=(x-1)x(x+1)(x+2)+1$,则

$$\begin{aligned}f(x)&=x^4+2x^3-x^2-2x+1\\&=(x^2+x-1)^2,\end{aligned}$$

于是 $f(x)$在 Q 上可约.

*8. 假如 $f(x)$在 Q 上可约,则存在 $f_1(x),f_2(x)\in Z[x]$使得

$$f(x)=f_1(x)f_2(x),\quad \deg f_i(x)<\deg f(x),\quad i=1,2.$$

于是
$$-1=f(a_i)=f_1(a_i)f_2(a_i),\quad i=1,2,\cdots,n.$$

从而
$$f_1(a_i)+f_2(a_i)=0,\quad i=1,2,\cdots,n.$$

由此推出,$f_1(x)+f_2(x)=0$,从而 $f(x)=-f_1^2(x)$.由于 $f(x)$的首项系数为1,因此 $f_1(x)$与 $f_2(x)$的首项系数同为1或同为-1,从而$-f_1^2(x)$的首项系数为-1,矛盾.

*9. 假如 $f(x)$在 Q 上可约,则存在 $f_1(x),f_2(x)\in Z[x]$使得

$$f(x)=f_1(x)f_2(x),\quad \deg f_i(x)<\deg f(x),\quad i=1,2.$$

于是
$$1=f(a_i)=f_1(a_i)f_2(a_i),\quad i=1,2,\cdots,n.$$

从而 $f_1(a_i)$与 $f_2(a_i)$同为1或同为-1,由于 $f(x)$没有实根,因此 $f_1(x),f_2(x)$也都没有实根,从而 $f_1(a_i)$全为1或全为-1,$i=1,2,\cdots,n$.不妨设 $f_1(a_1)=f_1(a_2)=\cdots=f_1(a_n)=1$.此时也有

$$f_2(a_1)=\cdots=f_2(a_n)=1.$$

情形1　$f_1(x)$与 $f_2(x)$中有一个多项式的次数小于 n,不妨设 $\deg f_1(x)<n$.由于

$f_1(a_i)=1, i=1,2,\cdots,n$,因此 $f_1(x)-1=0$,由此推出 $\deg f_1(x)=0$,$\deg f_2(x)=\deg f(x)$,矛盾.

情形 2 $\deg f_1(x)=\deg f_2(x)=n$.

由于 $f_1(a_i)-1=0, i=1,2,\cdots,n$.因此

$$f_1(x)-1=(x-a_1)(x-a_2)\cdots(x-a_n).$$

同理,

$$f_2(x)-1=(x-a_1)(x-a_2)\cdots(x-a_n).$$

从而

$$f(x)=f_1(x)f_2(x)=\prod_{i=1}^{n}(x-a_i)^2+2\prod_{i=1}^{n}(x-a_i)+1.$$

由此推出 $2\prod_{i=1}^{n}(x-a_i)=0$,矛盾.

习 题 7.9

1. (1) $f(x_1,x_2,x_3,x_4)=-x_1^3x_2+2x_2^4x_3x_4+5x_2x_3x_4+x_3^4x_4$;

(2) $f(x_1,x_2,x_3,x_4)=x_1^3-5x_1^2x_3x_4^2+3x_1x_2^2x_4-2x_2^3x_3+x_3^2$.

2. $f(x_1,x_2,x_3)=(x_1+x_2+x_3)(x_1^2+x_2^2+x_3^2-x_1x_2-x_1x_3-x_2x_3)$.

3. 假如 $f(x_1,x_2,\cdots,x_n)\neq 0$,由于 $g(x_1,x_2,\cdots,x_n)\neq 0$,因此 $f(x_1,x_2,\cdots,x_n)g(x_1,x_2,\cdots,x_n)\neq 0$,从而 fg 不是零函数.即存在 $c_1,c_2,\cdots,c_n\in K$,使得 $(fg)(c_1,c_2,\cdots,c_n)\neq 0$.从而 $f(c_1,c_2,\cdots,c_n)g(c_1,c_2,\cdots,c_n)\neq 0$.因此

$$f(c_1,c_2,\cdots,c_n)\neq 0,\text{且 } g(c_1,c_2,\cdots,c_n)\neq 0.$$

这与已知条件矛盾.因此 $f(x_1,x_2,\cdots,x_n)=0$.

*4. (1) 充分性显然.必要性.利用次数公式.

(2) 用反证法,注意利用次数公式.

习 题 7.10

1. 3 元对称多项式在含有一项 $x_1^3x_2^2$ 的同时,还应含有

$$x_1^3x_3^2,\quad x_2^3x_1^2,\quad x_2^3x_3^2,\quad x_3^3x_1^2,\quad x_3^3x_2^2.$$

由此看出 $f(x_1,x_2,x_3)$是 3 元对称多项式.

2. $x_1^3x_2+x_1^3x_3+x_2^3x_1+x_2^3x_3+x_3^3x_1+x_3^3x_2$.

3. (1) $\sigma_1^2\sigma_2-2\sigma_2^2-\sigma_1\sigma_3$;

(2) $\sigma_1^4-4\sigma_1^2\sigma_2+2\sigma_2^2+4\sigma_1\sigma_3$;

(3) $\sigma_1^3\sigma_3+\sigma_2^3-6\sigma_1\sigma_2\sigma_3+8\sigma_3^2$.

4. (1) $\sigma_1^3-3\sigma_1\sigma_2+3\sigma_3$;

(2) $n=3$ 时,$\sigma_2\sigma_3$;$n=4$ 时,$\sigma_2\sigma_3-3\sigma_1\sigma_4$;

$n\geq 5$ 时,$\sigma_2\sigma_3-3\sigma_1\sigma_4+5\sigma_5$.

5. $s_2=\sigma_1^2-2\sigma_2$, $s_3=\sigma_1^3-3\sigma_1\sigma_2+3\sigma_3$,

$s_4=\sigma_1^4-4\sigma_1^2\sigma_2+4\sigma_1\sigma_3+2\sigma_2^2$.

6. $D(f)=0$ 时,$f(x)$在 $\mathbf{C}$ 中有重根.由于 $\deg f(x)=3$,因此 $f(x)$至少有一个实根,

$f(x)$的另外两个根如果是共轭虚数,则 $f(x)$没有重根,矛盾.因此 $f(x)$的另外两个根也是实根.

如果 $f(x)$有三个不同的实根 r_1, r_2, r_3,则

$$D(f) = (r_2 - r_1)^2 (r_3 - r_1)^2 (r_3 - r_2)^2 > 0.$$

如果 $f(x)$有一个实根 r,一对共轭虚根 $c, \bar{c}$,设 $c = a + bi$,则

$$\begin{aligned} D(f) &= (c - r)^2 (\bar{c} - r)^2 (\bar{c} - c)^2 = |c - r|^4 (-2bi)^2. \\ &= -4b^2 |c - r|^4 < 0. \end{aligned}$$

因此,当 $D(f) > 0$ 时,$f(x)$的复根都是实数,且互不相同;当 $D(f) < 0$ 时,$f(x)$有一个实根,一对共轭虚根.

7. 设 $f(x)$的复根为 $c_1, c_2, \cdots, c_n$;则 $g(x)$的复根为 $a, c_1, c_2, \cdots, c_n$.于是

$$\begin{aligned} D(g) &= \prod_{1 \leq j < i \leq n} (c_i - c_j)^2 \prod_{k=1}^{n} (c_k - a)^2 \\ &= D(f) \prod_{k=1}^{n} (a - c_k)^2 = D(f) f(a)^2. \end{aligned}$$

8. 利用 Vieta 公式得

$$\sigma_1(c_1, c_2, \cdots, c_n) = \sigma_2(c_1, c_2, \cdots, c_n) = \cdots = \sigma_{n-1}(c_1, c_2, \cdots, c_n) = 0,$$

$$\sigma_n(c_1, c_2, \cdots, c_n) = (-1)^n a,$$

其中 $c_1, c_2, \cdots, c_n$ 是 $f(x) = x^n + a$ 的 n 个复根.利用牛顿公式求幂和;然后用本节公式(27)可求出

$$D(f) = (-1)^{\frac{1}{2}(n-1)(n-2)} a^{n-1} n^n.$$

9. $D(f) = -4a_1^3 a_3 + a_1^2 a_2^2 + 18 a_1 a_2 a_3 - 4a_2^3 - 27a_3^2$.

习 题 7.11

1. 设 $a \in F$ 且 $a \neq 0$.如果 $ab = 0$,则两边乘以 a^{-1} 得,$b = 0$,因此 F 没有非平凡的零因子.

2. Z_7 中,$\bar{1}^{-1} = \bar{1}$, $\bar{2}^{-1} = \bar{4}$, $\bar{3}^{-1} = \bar{5}$, $\bar{4}^{-1} = \bar{2}$, $\bar{5}^{-1} = \bar{3}$, $\bar{6}^{-1} = \bar{6}$.

*3. 直接按定义验证 F 是一个域.两个域同构是指它们作为环是同构的.令

$$\sigma: F \to C$$
$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} \longmapsto a + bi,$$

显然 σ 是双射,去验证 σ 保持加法和乘法,从而 σ 是 F 到 C 的一个同构映射.

4. 直接按定义验证 F 是一个域,以及 $\mathrm{char}\, F = 3$.

5. 用二项式定理.

6. $Z_2[x]$中一次多项式有:$x, x + 1$;二次不可约多项式只有 $x^2 + x + 1$(因为 0 和 1 都不是这个多项式的根,所以它没有一次因式.从而它在 Z_2 上不可约).这里我们把 $\bar{0}, \bar{1}$ 分别记成 0,1.以后同此约定.

7. 用$\tilde{f}(x)$表示 $f(x)$的各项系数模 2 以后得到的 $Z_2[x]$中的多项式:

$$\tilde{f}(x) = x^5 + x^2 + 1 = x^2(x + 1)(x^2 + x + 1) + 1,$$

用反证法可证$\tilde{f}(x)$在$Z_2[x]$中不可约.然后用反证法可证明$f(x)$在Q上不可约.

8. 把$f(x)$的各项系数模2得到$Z_2[x]$中的多项式:

$$\tilde{f}(x)=x^4+x^3+x^2+1=(x+1)(x^3+x+1).$$

容易说明x^3+x+1在Z_2上不可约.于是上式是$\tilde{f}(x)$在$Z_2[x]$中的唯一因式分解.假如$f(x)$在Q上可约,则

$$f(x)=f_1(x)f_2(x),\quad \deg f_i(x)<\deg f(x),\quad i=1,2.$$

把上式的每一个多项式的各项系数模2以后,得

$$\tilde{f}(x)=\tilde{f}_1(x)\tilde{f}_2(x),\quad \deg\tilde{f}_i(x)<\deg\tilde{f}(x),\quad i=1,2.$$

从$\tilde{f}(x)$在$Z_2[x]$中的唯一因式分解看出,$\tilde{f}_1(x)$与$\tilde{f}_2(x)$中必有一个是一次因式.从而$f_1(x)$与$f_2(x)$中必有一个是一次因式.由此推出$f(x)$有有理根.但是$f(x)$的有理根只可能是$\pm1,\pm5$.可以说明它们都不是$f(x)$的根.这个矛盾说明$f(x)$在Q上不可约.

第8章　线性空间

习　题　8.1

1. (1) 不是.提示:$2x^2+(-2x^2+x)=x$,不是2次多项式.

(2) 是.　　(3) 是.

2. (1) 线性相关.提示:利用二倍角公式$\cos 2x=2\cos^2 x-1$.

(2) 线性无关.提示:设$k_0+k_1\cos x+k_2\cos 2x+k_3\cos 3x=0$,选取$x$的4个恰当的值,代入上式得出$k_0,k_1,k_2,k_3$的4个方程的齐次线性方程组,说明它只有零解.

(3) 线性无关.提示:类似于第(2)小题的方法.

(4) 线性无关.提示:解法一.类似于第(2)小题方法;

解法二.设$k_1\sin x+k_2\cos x+k_3\sin^2 x+k_4\cos^2 x=0$,

在此式两边分别求1阶、2阶、3阶导数,然后令x取一个恰当的值,代入上述4个等式,得出k_1,k_2,k_3,k_4的4个方程的齐次线性方程组,说明它只有零解.

(5) 线性无关.　　(6) 线性无关.

3. 取定一个正实数$a\neq1$,则a是这个线性空间的一个基,从而维数是1.

4. $1,i$是一个基,维数是2.复数$z=a+bi$在基$1,i$下的坐标是(a,b).

5. 1是一个基,从而维数是1.

6. $E_{11},E_{22},\cdots,E_{nn},E_{12}+E_{21},\cdots,E_{1n}+E_{n1},\cdots,E_{n-1,n}+E_{n,n-1}$是$V_1$的一个基,从而$V_1$的维数是$\dfrac{n(n+1)}{2}$.

7. V_2的一个基是:

$$E_{12}-E_{21},\cdots,E_{1n}-E_{n1},E_{23}-E_{32},\cdots,E_{2n}-E_{n2},\cdots,E_{n-1,n}-E_{n,n-1}.$$

V_1的维数是$\dfrac{n(n-1)}{2}$.

8. W 的一个基是:

$$E_{11}, E_{12}, \cdots, E_{1n}, E_{22}, E_{23}, \cdots, E_{2n}, \cdots, E_{n-1,n-1}, E_{n-1,n}, E_{nn}.$$

W 的维数是$\frac{n(n+1)}{2}$.

9. 提示:设 $A=(\alpha_1,\alpha_2,\alpha_3)$, $B=(\beta_1,\beta_2,\beta_3)$,由于

$$(\beta_1,\beta_2,\beta_3)=(\alpha_1,\alpha_2,\alpha_3)P,$$

因此 $B=AP$.解这个矩阵方程可求得

$$P=\begin{pmatrix}0 & 1 & 1\\ -1 & -3 & -2\\ 2 & 4 & 4\end{pmatrix}.$$

由于 $\alpha=(\beta_1,\beta_2,\beta_3)Y$,因此 $BY=\alpha$.解之,得 $Y=(1,0,2)'$.进而

$$X=PY=(2,-5,10)'.$$

10. 提示:在 V 中取一个基 $\delta_1,\delta_2,\cdots,\delta_n$,由已知条件得,$\delta_1,\delta_2,\cdots,\delta_n$ 可由 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 线性表出,从而 $\text{rank}\{\delta_1,\delta_2,\cdots,\delta_n\}\leq\text{rank}\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$,由此求出 $\text{rank}\{\alpha_1,\alpha_2,\cdots,\alpha_n\}$,进而判断 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 必线性无关.

11. 任给 $f\in F^X$, f 被它的所有函数值 $f(x_1), f(x_2),\cdots, f(x_n)$决定.构造 n 个函数如下:

$$f_i(x_j)=\delta_{ij}, 1\leq j\leq n, i=1,2,\cdots,n.$$

去证 f 可以由 $f_1,f_2,\cdots,f_n$ 线性表出;然后证 $f_1,f_2,\cdots,f_n$ 线性无关.从而 $f_1,f_2,\cdots,f_n$ 是 F^X 的一个基,于是 $\dim F^X=n$.函数 f 在这个基下的坐标为$(f(x_1), f(x_2),\cdots, f(x_n))'$.

12. (1) 按照线性空间的定义去验证;

(2) V 的一个基为

$$\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \begin{pmatrix}0 & i\\ -i & 0\end{pmatrix}, \quad \dim V=3;$$

(3) 在上述基下的坐标为

$$(x_1, x_2, x_3)'.$$

习 题 8.2

1. (1) 是; (2) 不是; (3) 不是.

2. 提示:去证 $C(A)$非空集,对加法封闭,对数乘封闭.

3. 提示:利用习题 4.2 的第 1 题的结论.$C(A)$的一个基为 $E_{11}, E_{22},\cdots, E_{nn}$;于是 $\dim C(A)=n$.

4. 设 $X=(x_{ij})\in C(A)$,则 $AX=XA$.由此得出一个 9 元齐次线性方程组.容易求出 x_{11}, x_{12}, x_{13}是自由未知量,从而可求出 $C(A)$的一个基为

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 2 & 0\\ 1 & 4 & -1\\ 2 & 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0 & 1\\ 1 & 0 & 0\\ 4 & -2 & 1\end{pmatrix}.$$

5. 提示:去证 $\alpha\in\langle\beta,\gamma\rangle$, $\beta\in\langle\alpha,\gamma\rangle$.

6. 维数为 3,一个基为 $\alpha_1,\alpha_2,\alpha_3$.

7. 提示:设 A 的列向量组 $A_1,A_2,\cdots,A_s$ 的一个极大线性无关组是 $A_{j_1},\cdots,A_{j_r}$,去证 $\beta_{j_1},\cdots,\beta_{j_r}$ 是向量组 $\beta_1,\beta_2,\cdots,\beta_s$ 的一个极大线性无关组.

*8. 提示:由于符号差 $s>0$,因此 $p>(n-p)$.显然

$$\alpha_1=(1,0,\cdots,0,1,0,0,\cdots,0,0),$$

第 $p+1$ 位

$$\alpha_2=(0,1,\cdots,0,0,1,0,\cdots,0,0),$$

…………

$$\alpha_{n-p}=(0,0,\cdots,0,1,\cdots,0,0,0,0,\cdots,0,1)$$

都是 $f(x_1,x_2,\cdots,x_n)=0$ 的解,它们的任意一个线性组合 $k_1\alpha_1+k_2\alpha_2+\cdots+k_{n-p}\alpha_{n-p}$ 也是 $f(x_1,x_2,\cdots,x_n)=0$ 的解.从而 $\langle\alpha_1,\alpha_2,\cdots,\alpha_{n-p}\rangle\subseteq W$.

9. 提示:由于 $V_1\neq V$,因此存在 $\alpha\notin V_1$.如果 $\alpha\notin V_2$,则 $\alpha\notin V_1\cup V_2$.如果 $\alpha\in V_2$,则存在 $\beta\notin V_2$.若 $\beta\notin V_1$,则 $\beta\notin V_1\cup V_2$.若 $\beta\in V_1$,去证 $\alpha+\beta\notin V_1\cup V_2$.

*10. 提示:对 s 用数学归纳法.当 $s=2$ 时,第 9 题已证明 $V_1\cup V_2\neq V$.假设命题对于 $s-1$ 的情形成立,考虑 s 的情形.据归纳假设,V 中存在向量 $\alpha\notin V_1\cup V_2\cup\cdots\cup V_{s-1}$.如果 $\alpha\notin V_s$,则 $\alpha\notin V_1\cup V_2\cup\cdots\cup V_{s-1}\cup V_s$.如果 $\alpha\in V_s$,设 $\beta\notin V_s$.考虑集合 $W=\{k\alpha+\beta\mid k\in F\}$.易看出,对任意 $k\in F$,有 $k\alpha+\beta\notin V_s$.再证对于每一个 $V_i(i=1,2,\cdots,s-1)$,W 中至多有一个向量属于 V_i.从而 W 中至少有一个向量 γ 不属于 $V_1\cup V_2\cup\cdots\cup V_{s-1}$.于是 $\gamma\notin V_1\cup V_2\cup\cdots\cup V_{s-1}\cup V_s$.

11. V_1+V_2 的一个基是 $\alpha_1,\alpha_2,\beta_1$,$\dim(V_1+V_2)=3$;

$V_1\cap V_2$ 的一个基是 $(0,1,1,-1)'$,$\dim(V_1\cap V_2)=1$.

12. V_1+V_2 的一个基是 $\alpha_1,\alpha_2,\beta_1$,$\dim(V_1+V_2)=3$;

$V_1\cap V_2$ 的一个基是 $(5,-1,5,2)'$,$\dim(V_1\cap V_2)=1$.

13. 提示:先证 $V=V_1+V_2$,关键是证 V 中任一向量 $\alpha=(a_1,a_2,\cdots,a_n)'$ 能表示成 $\alpha_1+\alpha_2$,其中 $\alpha_1\in V_1$,$\alpha_2\in V_2$.可令

$$\alpha_2=\left(\frac{1}{n}\sum_{i=1}^{n}a_i,\frac{1}{n}\sum_{i=1}^{n}a_i,\cdots,\frac{1}{n}\sum_{i=1}^{n}a_n\right)'.$$

再去证 $V_1\cap V_2=0$.

14. 提示:取 V 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$.先说明

$$V=\langle\alpha_1\rangle+\langle\alpha_2\rangle+\cdots+\langle\alpha_n\rangle,$$

再说明上式右边的和是直和.

15. (1) 提示:去证 $M_n^0(K)$ 非空集,对加法封闭,对数乘封闭;

(2) 提示:先证 $M_n(K)=\langle I\rangle+M_n^0(K)$,关键是证任一 n 级矩阵 $A=(a_{ij})$ 能表示成 A_1+A_2,其中 $A_1\in\langle I\rangle$,$A_2\in M_n^0(K)$.再证 $\langle I\rangle\cap M_n^0(K)=0$.

*16. 提示:充分性.利用第五章§6 的定理 5.

必要性.设 A 可对角化,据第五章§6 的定理 5 得,

$$\dim V_{\lambda_1} + \dim V_{\lambda_2} + \cdots + \dim V_{\lambda_s} = n.$$

在 $V_{\lambda_1}, V_{\lambda_2}, \cdots, V_{\lambda_s}$ 中各取一个基,它们合起来是 n 个线性无关的向量,成为 K^n 的一个基.因此 $K^n = V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_s}$.从而 $V_{\lambda_1} + V_{\lambda_2} + \cdots + V_{\lambda_s}$ 是直和.

17. (1) 提示:只要证 W_1　W, W_2　W.

(2) 提示:由于 $(f_1(x), f_2(x)) = 1$,因此有 $u_i(x) \in K[x]$ $(i = 1, 2)$,使得 $u_1(x) f_1(x) + u_2(x) f_2(x) = 1$.不定元 x 用 A 代入得, $u_1(A) f_1(A) + u_2(A) f_2(A) = I$,利用此等式去证 W　$W_1 + W_2$,以及 $W_1 \cap W_2 = 0$.

习　题　8.3

1. 提示:对于任意实数 x,令 $\sigma: x \longmapsto 2^x$,去证 σ 是 R 到 R^+ 的一个同构映射.

2. 提示:因为 $\dim M_{s\times n}(F) = sn = \dim F^{sn}$,所以 $M_{s\times n}(F) \cong F^{sn}$,一个同构映射 σ 是:对于 $A = (a_{ij})$, $\sigma(A) = (a_{11}, a_{12}, \cdots, a_{1n}, a_{21}, a_{22}, \cdots, a_{2n}, \cdots, a_{s1}, a_{s2}, \cdots, a_{sn})$.

3. 提示:因为 $\dim F[x]_n = n = \dim F^n$,所以 $F[x]_n \cong F^n$.一个同构映射 σ 是:对于 $f(x) = a_0 + a_1 x + \cdots + a_{n-1} x^{n-1}$, $\sigma(f(x)) = (a_0, a_1, \cdots, a_{n-1})$.

4. 提示:令 $U = \langle \alpha_1, \alpha_2, \alpha_3, \alpha_4 \rangle$,则 $\dim U = 4$,设 σ 是 U 到 F^4 的一个同构映射,它把 U 中每一个向量对应到该向量在基 $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ 下的坐标.求出 $\sigma(\alpha_1 + \alpha_2)$, $\sigma(\alpha_2 + \alpha_3)$, $\sigma(\alpha_3 + \alpha_4)$, $\sigma(\alpha_4 + \alpha_1)$,把以它们为列向量的矩阵经过初等行变换化成阶梯形矩阵,便可看出 $\sigma(\alpha_1 + \alpha_2)$, $\sigma(\alpha_2 + \alpha_3)$, $\sigma(\alpha_3 + \alpha_4)$, $\sigma(\alpha_4 + \alpha_1)$ 是否线性无关,进而可求出它们生成的 F^4 的子空间的一个基和维数,由此可求出 W 的一个基和维数:基为 $\alpha_1 + \alpha_2, \alpha_2 + \alpha_3, \alpha_3 + \alpha_4$,维数为 3.

5. (1) 提示:显然 L 非空集,易看出 L 对于加法、数量乘法都封闭. L 的一个基是 $E_{11} + E_{22}$, $E_{12} - E_{21}$; $\dim L = 2$.

(2) 提示:因为 $\dim_R C = 2 = \dim L$,所以 $C \cong L$.一个同构映射 σ 是:

$$\sigma(a + bi) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

*6. (1) 提示:按照线性空间的定义去验证;

(2) 一个基为 $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \begin{pmatrix} 0, & i \\ i & 0 \end{pmatrix}$, $\dim H = 4$;

(3) 提示:注意 $\dim H = 4 = \dim R^4$;一个同构映射 τ 是:设 $\alpha = a + bi, \beta = c + di$,则

$$\tau \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} = (a, c, b, d).$$

习　题　8.4

1. 提示:去证 $w \longmapsto w + U$ 是 W 到 V　U 的一个同构映射.

2. 提示: V　U 由所有与 U 平行或重合的直线组成.在 U 中取一个基 u_1,把它扩充成 V

的一个基: u_1, α_1, α_2,去证 $V\ U$ 中任一元素 $\alpha + U$ 可以由 $\alpha_1 + U, \alpha_2 + U$ 线性表出,再证 $\alpha_1 + U, \alpha_2 + U$ 线性无关.从而 $\alpha_1 + U, \alpha_2 + U$ 是 $V\ U$ 的一个基,于是 $\dim V\ U = 2$.维数的另一求法:

$$\dim V\ U = \dim V - \dim U = 3 - 1 = 2.$$

3. 提示:任取 Ω_2 的一个元素 $W\ U$,则 $\dim W = \dim U + \dim W\ U = n - 1$.从而 W 是 V 的 $n-1$ 维子空间,且 W 包含 U.于是 $W \in \Omega_1$.由此可证 σ 是满射.假设 Ω_2 中, $W_1\ U = W_2\ U$,分别在 $W_1\ U$, $W_2\ U$ 中取一个基: $\alpha_1 + U, \cdots, \alpha_{n-3} + U$; $\beta_1 + U, \cdots, \beta_{n-3} + U$.在 U 中取一个基 u_1, u_2,则易看出,

$$W_1 = \langle \alpha_1, \cdots, \alpha_{n-3}, u_1, u_2 \rangle, \quad W_2 = \langle \beta_1, \cdots, \beta_{n-3}, u_1, u_2 \rangle.$$

由于 $W_1\ U = W_2\ U$,因此 $\alpha_i + U \in W_2\ U, i = 1, 2, \cdots, n-3$.由此可证 $\langle \alpha_1, \cdots, \alpha_{n-3}, u_1, u_2 \rangle = \langle \beta_1, \cdots, \beta_{n-3}, u_1, u_2 \rangle$.即 $W_1 = W_2$.从而 σ 是单射.

4. 提示:令

$$\sigma: \quad U + W\ W \quad U\ U \cap W$$
$$(u + w) + W \longmapsto u + U \cap W,$$

去证 σ 是映射、单射、满射,且保持加法、纯量乘法运算.

第 9 章　线性映射

习　题　9.1

1. (1) 不是.(2) 是.　2. (1) 是;(2) 是.　3. 是.　4. 是.

*5. 是(提示:利用多元多项式环的通用性质).

*6. 是(提示:利用在特征为 2 的域中,有 $(x + y)^2 = x^2 + y^2$).

7. (1) 提示:按线性变换的定义验证.　(2) 提示:直接计算.

8. 提示:必要性.利用 V 上的线性变换 A 可逆当且仅当 A 是 V 到自身的一个同构映射.充分性.去证 A 是单射,满射.

9. 提示:设 $k_0 \alpha + k_1 A\alpha + k_2 A^2 \alpha + \cdots + k_{m-1} A^{m-1} \alpha = 0$,此式两边用 A^{m-1} 作用,得 $k_0 A^{m-1} \alpha = 0$,由此得出 $k_0 = 0$.从而有 $k_1 A\alpha + k_2 A^2 \alpha + \cdots + k_{m-1} A^{m-1} \alpha = 0$.

10. 提示:对 k 用数学归纳法.

11. (1) 提示:必要性.从 $(A + B)^2 = A + B$ 可得出, $AB + BA = 0$,此式两边左乘 A,且用 $AB = -BA$ 代入,可推出 $AB = BA$.

(2) 提示:去证 $(A + B - AB)^2 = A + B - AB$.

习　题　9.2

1. 提示:由 $\operatorname{Ker} A, \operatorname{Im} A$ 的定义立即得到.

2. 提示:先证 $V = \operatorname{Ker} A + \operatorname{Im} A$,为此任取 $\alpha \in V$,去证 $\alpha - A\alpha \in \operatorname{Ker} A$.再证 $\operatorname{Ker} A \cap \operatorname{Im} A$

$=0$.

3. 提示:考虑 $\mathcal{B}$ 在 $\mathcal{A}V$ 上的限制:$\mathcal{B}|\mathcal{A}V$,它的象是 $\mathcal{B}(\mathcal{A}V)=(\mathcal{B}\mathcal{A})V$.计算 $\dim \operatorname{Ker}(\mathcal{B}\mathcal{A})=\dim V-\dim(\mathcal{B}\mathcal{A})V=\cdots$.

4. 提示:利用本节定理 3 和第 3 题的结论.

5. (1) 提示:必要性.利用第 2 题的结论以及§1 的定理 2.充分性.任取 $\delta\in \operatorname{Im}\mathcal{A}$,则存在 $\alpha\in V$ 使得 $\mathcal{A}\alpha=\delta$.利用 $\mathcal{B}\mathcal{A}=\mathcal{A}$ 可推出 $\delta\in \operatorname{Im}\mathcal{B}$.

(2) 提示:必要性.利用第 2 题的结论得,对于 $\alpha\in V$,有 $\alpha=\alpha_1+\alpha_2$,其中 $\alpha_1\in \operatorname{Ker}\mathcal{B}$,$\alpha_2\in \operatorname{Im}\mathcal{B}$,从而 $\mathcal{B}\alpha=\mathcal{B}\alpha_2=\alpha_2$.因此 $\mathcal{B}\alpha=\alpha-\alpha_1$,此式两边左乘 $\mathcal{A}$,可推出 $\mathcal{A}\mathcal{B}=\mathcal{A}$.充分性.任取 $\delta\in \operatorname{Ker}\mathcal{A}$,利用 $\mathcal{B}\mathcal{A}=\mathcal{B}$ 可证 $\delta\in \operatorname{Ker}\mathcal{B}$.

6. 提示:取 $U=\operatorname{Ker}\mathcal{A}$,由于 V 是有限维的,因此存在 U 的补空间 W.设 $\dim V=n$,$\dim W=r$.去证 $\dim \mathcal{A}V=r$.从而 $W\cong \mathcal{A}V$.取 $\mathcal{M}=\mathcal{A}V$.

*7. 提示:当 $i\neq j$ 时,由于 $\mathcal{A}_i\neq \mathcal{A}_j$,因此 $\mathcal{A}_i-\mathcal{A}_j\neq 0$,从而 $\operatorname{Ker}(\mathcal{A}_i-\mathcal{A}_j)\neq V$.然后利用习题 8.2 的第 10 题的结论.

习　题　9.3

1. $\begin{pmatrix}1&2&0\\0&-1&1\\0&1&1\end{pmatrix}$.　2. $\begin{pmatrix}a&b\\-b&a\end{pmatrix}$.　3. $\begin{pmatrix}a&0&b&0\\0&a&0&b\\c&0&d&0\\0&c&0&d\end{pmatrix}$.

4. $\begin{pmatrix}a&c&0&0\\b&d&0&0\\0&0&a&c\\0&0&b&d\end{pmatrix}$.

5. 提示:据习题 9.1 的第 9 题结论得,$\alpha,\mathcal{A}\alpha,\cdots,\mathcal{A}^{n-1}\alpha$ 线性无关,考虑 $\mathcal{A}$ 在基 $\mathcal{A}^{n-1}\alpha,\mathcal{A}^{n-2}\alpha,\cdots,\mathcal{A}\alpha,\alpha$ 下的矩阵.

6. 提示:线性变换与它的矩阵的对应是保持乘法的.

7. (1) 提示:由于 $\dim \operatorname{Hom}(V,V)=(\dim V)^2=n^2$,因此 V 上的 n^2+1 个线性变换必定线性相关,从而有 F 的不全为 0 的元素 $k_0,k_1,\cdots,k_{n^2}$,使得

$$k_0\mathcal{I}+k_1\mathcal{A}+k_2\mathcal{A}^2+\cdots+k_{n^2}\mathcal{A}^{n^2}=0.$$

令 $f(x)=k_0+k_1x+k_2x^2+\cdots+k_{n^2}x^{n^2}$,不定元 x 用 $\mathcal{A}$ 代入.

(2) 提示:存在 $u(x),v(x)\in F[x]$,使得

$$u(x)f(x)+v(x)g(x)=d(x),$$

不定元 x 用 $\mathcal{A}$ 代入.

(3) 提示:充分性.设 $f(x)=a_0+a_1x+\cdots+a_mx^m\ (a_0\neq 0)$ 使得 $f(\mathcal{A})=0$.不定元 x 用 $\mathcal{A}$ 代入,然后根据可逆变换的定义去证.必要性.据第(1)小题,可以取一个次数最低的非零多项式 $f(x)$ 使得 $f(\mathcal{A})=0$,去证 $f(x)$ 的常数项一定不为 0.

8. 提示:由已知条件得,

$$(\mathcal{A}\alpha_1, \mathcal{A}\alpha_2, \cdots, \mathcal{A}\alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)A.$$

利用第 8 章 § 3 例 1 的结论;然后用 rank A 的定义.

*9. 提示:设 $\dim V = n$,则 $\operatorname{Ker}\mathcal{A}$ 在 V 中有补空间 W,因此 $V = \operatorname{Ker}\mathcal{A} \oplus W$.在 W 中取一个基 $\alpha_1, \cdots, \alpha_r$,在 $\operatorname{Ker}\mathcal{A}$ 中取一个基:$\alpha_{r+1}, \cdots, \alpha_n$,于是可得 V 的一个基.然后利用习题 9.2 的第 6 题的证明过程知道,$V' = \mathcal{A}V \oplus N$.易知 $\mathcal{A}\alpha_1, \cdots, \mathcal{A}\alpha_r$ 是 $\mathcal{A}V$ 的一个基,再在 N 中取一个基,从而可得 V' 的一个基.考虑 $\mathcal{A}$ 在 V 的上述基和 V' 的这个基下的矩阵.

10. 提示:先求基 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 到基 η_1, η_2, η_3 的过渡矩阵 S,然后求 S^{-1},最后求 $B = S^{-1}AS$,得出

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

11. 提示:记 $C = (\alpha_1, \alpha_2, \alpha_3)$, $T = (\eta_1, \eta_2, \eta_3)$.设 $(\eta_1, \eta_2, \eta_3) = (\alpha_1, \alpha_2, \alpha_3)S$,则 $S = C^{-1}T$.从而 $B = S^{-1}AS = (C^{-1}T)^{-1}A(C^{-1}T)$,计算得

$$B = \begin{pmatrix} 1 & 2 & 2 \\ 3 & -1 & -2 \\ 2 & 3 & 1 \end{pmatrix}.$$

12. (1) 提示:$(\alpha_2, \alpha_3, \alpha_1) = (\alpha_1, \alpha_2, \alpha_3)\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$,

过渡矩阵 S 是正交矩阵,于是 $S^{-1} = S'$.计算 $S^{-1}AS$,得

$$\begin{pmatrix} a_{22} & a_{23} & a_{21} \\ a_{32} & a_{33} & a_{31} \\ a_{12} & a_{13} & a_{11} \end{pmatrix};$$

(2) 提示:

$$(k\alpha_1, \alpha_2, \alpha_3) = (\alpha_1, \alpha_2, \alpha_3)\begin{pmatrix} k & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

过渡矩阵 S 是对角矩阵,计算 $S^{-1}AS$,得

$$\begin{pmatrix} a_{11} & k^{-1}a_{12} & k^{-1}a_{13} \\ ka_{21} & a_{22} & a_{23} \\ ka_{31} & a_{32} & a_{33} \end{pmatrix};$$

(3) 提示:$(\alpha_1, \alpha_1 + \alpha_2, \alpha_3) = (\alpha_1, \alpha_2, \alpha_3)\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

计算 $S^{-1}AS$,得

$$\begin{pmatrix} a_{11} - a_{21} & a_{11} + a_{12} - a_{21} - a_{22} & a_{13} - a_{23} \\ a_{21} & a_{21} + a_{22} & a_{23} \\ a_{31} & a_{31} + a_{32} & a_{33} \end{pmatrix}.$$

*13. 提示:设 $\mathcal{A}$ 在 V 的一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵为 A.任取域 F 上一个 n 级可逆矩阵 S,令

$$(\eta_1,\eta_2,\cdots,\eta_n)=(\alpha_1,\alpha_2,\cdots,\alpha_n)S.$$

则 $\eta_1,\eta_2,\cdots,\eta_n$ 也是 V 的一个基,且 $\mathcal{A}$ 在基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的矩阵 $B=S^{-1}AS$.由已知条件得,$S^{-1}AS=B=A$.从而 $AS=SA$.取 $S=\mathrm{diag}\{1,2,\cdots,n\}$,从 $AS=SA$,利用上册第 4 章习题 4.2 的第 1 题可得出,A 必为对角矩阵.取 $S=I+E_{1j}$(其中 $2\leq j\leq n$),从 $AS=SA$ 可推出 A 必为数量矩阵.

14. (1)

$$\begin{matrix} 2 & -3 & 3 & 2 \\ -\frac{2}{3} & -\frac{4}{3} & \frac{10}{3} & \frac{10}{3} \\ \frac{8}{3} & -\frac{16}{3} & \frac{40}{3} & \frac{40}{3} \\ 0 & 1 & -7 & -8 \end{matrix}.$$

(2) 提示:设 α 在基 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 下的坐标为 X,则

$$\alpha\in \mathrm{Ker}\,\mathcal{A} \qquad \mathcal{A}\alpha=0 \qquad AX=0$$

去求齐次线性方程组 $AZ=0$ 的一个基础解系,以它们为坐标的向量就组成 $\mathrm{Ker}\mathcal{A}$ 的一个基.

设 β 在基 $\alpha_1,\alpha_2,\alpha_3,\alpha_4$ 下的坐标为 Y,则

$$\beta\in \mathcal{A}V \qquad \beta\in\langle\mathcal{A}\alpha_1,\mathcal{A}\alpha_2,\mathcal{A}\alpha_3,\mathcal{A}\alpha_4\rangle$$

$$Y\in\langle A_1,A_2,A_3,A_4\rangle$$

去求 A 的列空间 $\langle A_1,A_2,A_3,A_4\rangle$ 的一个基,以它们为坐标的向量就组成 $\mathcal{A}V$ 的一个基.

$\mathrm{Ker}\,\mathcal{A}$ 的一个基是:$4\alpha_1+3\alpha_2-2\alpha_3,\alpha_1+2\alpha_2-\alpha_4$;

$\mathcal{A}V$ 的一个基是:$\alpha_1-\alpha_2+\alpha_3+2\alpha_4,2\alpha_2+2\alpha_3-2\alpha_4$.

(3) 提示:$\mathrm{Ker}\,\mathcal{A}$ 的一个基可以取第(2)小题中求出的基,以它们的坐标为矩阵的第 1、2 列,添上第 3、4 列,使 4 级矩阵的行列式不为 0.以第 3、4 列为坐标的向量添到 $\mathrm{Ker}\,\mathcal{A}$ 的一个基上,即得 V 的一个基.

(4) 提示:可取第(2)小题中求出的 $\mathrm{Im}\,\mathcal{A}$ 的一个基,然后类似于第(3)小题的方法,可扩充成 V 的一个基.

习　题　9.4

1. (1) $\mathcal{A}$ 的全部特征值是 1(二重),10.

$\mathcal{A}$ 的属于特征值 1 的全部特征向量是

$$\{k_1(-2\alpha_1+\alpha_2)+k_2(2\alpha_1+\alpha_3)\mid k_1,k_2\in K,\text{且 } k_1,k_2 \text{ 不全为 } 0\};$$

$\mathcal{A}$ 的属于特征值 10 的全部特征向量是

$$\{k(\alpha_1+2\alpha_2-2\alpha_3)\mid k\in K,\text{且 } k\neq 0\};$$

(2) $\mathcal{A}$ 的全部特征值是 1,3(二重).

$\mathcal{A}$ 的属于 1 的全部特征向量是

$$\{k(-2\alpha_1+\alpha_3)\mid k\in K,\text{且 } k\neq 0\};$$

A 的属于 3 的全部特征向量是

$$\{k(\alpha_1-\alpha_2+2\alpha_3)\mid k\in K,\text{且 } k\neq 0\}.$$

2. (1) A 可对角化, A 的标准形是

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&10\end{pmatrix};$$

(2) A 不可以对角化.

3. (1) A 的全部特征值是 1(二重),0(二重).

A 的属于 1 的全部特征向量是

$$\{k_1(\alpha_1+\alpha_3)+k_2\alpha_4\mid k_1,k_2\in K,\text{且 } k_1,k_2 \text{ 不全为 } 0\};$$

A 的属于 0 的全部特征向量是

$$\{l_1\alpha_2+l_2\alpha_3\mid l_1,l_2\in K,\text{且 } l_1,l_2 \text{ 不全为 } 0\}.$$

(2) A 在 V 的一个基 $\alpha_1+\alpha_3,\alpha_4,\alpha_2,\alpha_3$ 下的矩阵为

$$\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}.$$

4. 提示:设 λ_1,λ_2 是 A 的不同特征值, ξ_1,ξ_2 分别是 A 的属于 λ_1,λ_2 的一个特征向量.假设 $k_1\xi_1+k_2\xi_2=0$,去证 $k_1=k_2=0$.

5. 提示:用反证法,并且利用第 4 题的结论.

6. 提示:假如 A 有两个不同的特征值 λ_1,λ_2,利用第 5 题的结论,将与已知条件矛盾.

7. (1) 提示:用反证法,将与 A 是单射矛盾.

(2) 提示:考虑 $A\xi=\lambda\xi$ 两边的向量在 A^{-1} 下的象.

8. 提示: A 可对角化当且仅当 V 中存在由 A 的特征向量组成的一个基: $\xi_{i1},\xi_{i2},\cdots,\xi_{ir_i}$, $i=1,2,\cdots,s$;其中 $\xi_{i1},\xi_{i2},\cdots,\xi_{ir_i}$ 是 A 的属于 λ_i 的线性无关的特征向量.

*9. 提示:易知 A 可对角化,于是 V 中存在一个基 $\alpha_1,\alpha_2,\cdots,\alpha_n$,使得 A 在此基下的矩阵 $A=\text{diag}\{\lambda_1,\lambda_2,\cdots,\lambda_n\}$.利用 B 与 A 可交换,说明 B 在此基下的矩阵 B 必为对角矩阵.设 $B=c_0I+c_1A+\cdots+c_{n-1}A^{n-1}$,则 $B=c_0I+c_1A+\cdots+c_{n-1}A^{n-1}$.说明 $(c_0,c_1,\cdots,c_{n-1})$ 存在且唯一.

习 题 9.5

1. 提示:计算 $A(\alpha_1+2\alpha_2)$, $A(\alpha_2+\alpha_3+2\alpha_4)$.

2. (1) 提示:先说明 $A|W$ 是单射,然后利用 $A|W$ 是单射的结论说明 $A|W$ 是满射.

(2) 提示:用第(1)小题的结论,根据定义去证 W 是 A^{-1} 的不变子空间.任取 $\beta\in W$,设 $(A|W)^{-1}\beta=\gamma$,则 $\gamma\in W$.去证 $\gamma=(A^{-1}|W)\beta$.

3. 提示:取 A 的一个特征值 λ_0,则 V_{λ_0} 是 B-子空间.取 $B|V_{\lambda_0}$ 的一个特征值 μ_0,则存在 $\xi\in V_{\lambda_0}$,使得 $B\xi=(B|V_{\lambda_0})\xi=\mu_0\xi$.

4. (1) 提示:如果 $\alpha_n\in W$,则由于 W 是 $\mathcal{A}$-子空间,因此 $\mathcal{A}\alpha_n\in W$,即 $\alpha_{n-1}+a\alpha_n\in W$,从而 $\alpha_{n-1}\in W$.同理有 $\mathcal{A}\alpha_{n-1}\in W$,依次下去.

(2) 提示:设 W 是 $\mathcal{A}$ 的非零不变子空间,取 $\beta\in W$,且 $\beta\neq 0$.设 $\beta=k_1\alpha_1+\cdots+k_s\alpha_s$,其中 $k_s\neq 0$.由 $\mathcal{A}\beta\in W$ 可推出 $k_2\alpha_1+\cdots+k_s\alpha_{s-1}\in W$.继续用 $\mathcal{A}$ 作用下去.

(3) 提示:利用第(2)小题的结论.

(4) 提示:设 W 是非零的 $\mathcal{A}$-子空间,则 $\alpha_1\in W$.设 $\dim W=m$,在 W 中取一个基 α_1, $\beta_2,\cdots,\beta_m$.设

$$\beta_2=k_{21}\alpha_1+k_{22}\alpha_2+\cdots+k_{2s}\alpha_s,$$
$$\cdots\cdots\cdots\cdots$$
$$\beta_m=k_{m1}\alpha_1+k_{m2}\alpha_2+\cdots+k_{ms}\alpha_s,$$

其中 $k_{2s},\cdots,k_{ms}$不全为零,不妨设 $k_{2s}\neq 0$.由于 $\alpha_1,\beta_2,\cdots,\beta_m$ 可由$\alpha_1,\alpha_2,\cdots,\alpha_s$ 线性表出,因此 $m\leq s$.由于 $\mathcal{A}\beta_2\in W$,因此 $k_{21}(a\alpha_1)+k_{22}(\alpha_1+a\alpha_2)+\cdots+k_{2s}(\alpha_{s-1}+a\alpha_s)\in W$.由此推出,$k_{23}\alpha_2+\cdots+k_{2s}\alpha_{s-1}\in W$.继续用 $\mathcal{A}$ 作用下去,可推出 $k_{2s}\alpha_2\in W$,从而 $\alpha_2\in W$.同理可证 $\alpha_3\in W,\cdots,\alpha_s\in W$.于是 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 可由$\alpha_1,\beta_2,\cdots,\beta_m$ 线性表出,因此 $s\leq m$.从而 $s=m$.因此 $\alpha_1,\alpha_2,\cdots,\alpha_m$ 是 W 的一个基.即, $W=\langle\alpha_1,\alpha_2,\cdots,\alpha_m\rangle$.从而 $\mathcal{A}$ 的所有不变子空间共 $n+1$ 个,它们是:

$$\{0\},\langle\alpha_1\rangle,\langle\alpha_1,\alpha_2\rangle,\cdots,\langle\alpha_1,\alpha_2,\cdots,\alpha_n\rangle.$$

5. 提示:若 $\mathcal{A}$ 有特征值,取 $\mathcal{A}$ 的一个特征值λ_0,设 ξ是 $\mathcal{A}$ 的属于λ_0 的一个特征向量,则易证$\langle\xi\rangle$是 $\mathcal{A}$ 的1维不变子空间.若 $\mathcal{A}$ 没有特征值,设 $\mathcal{A}$ 在 V 的一个基$\alpha_1,\alpha_2,\cdots,\alpha_n$ 下的矩阵为 A.把 A 看成复数域上的矩阵,则 A 有特征值.由于 A 的特征多项式是实系数多项式,因此复矩阵 A 的特征值共轭成对出现.设 $z,\bar{z}$是 A 的一对特征值,其中 $z=a+bi$, $a,b\in\mathbf{R}$.设 $X=X_1+\mathrm{i}X_2$ 是 A 的属于 z的一个特征向量,则$\overline{X}=X_1-\mathrm{i}X_2$ 是 A 的属于$\bar{z}$的一个特征向量,其中 $X_1,X_2\in\mathbf{R}^n$.从 $AX=zX$, $A\overline{X}=\bar{z}\overline{X}$可推出,

$$AX_1=aX_1-bX_2,\quad AX_2=bX_1+aX_2.$$

令 $\beta_1=(\alpha_1,\alpha_2,\cdots,\alpha_n)X_1$,$\beta_2=(\alpha_1,\alpha_2,\cdots,\alpha_n)X_2$.去证$\langle\beta_1,\beta_2\rangle$是 $\mathcal{A}$ 的不变子空间;再去证β_1,β_2 线性无关.

6. 提示:用反证法,注意 $\mathcal{A}$ 没有特征向量.

*7. 提示: V_{λ_0}是 $\mathcal{A}$-子空间,且 $V_{\lambda_0}\neq 0$.若 $\mathcal{A}$ 是数乘变换,则结论显然成立.下面设 $\mathcal{A}$ 不是数乘变换,于是 $V_{\lambda_0}\neq V$.取 V_{λ_0} 的一个基 $\alpha_1,\cdots,\alpha_r$,把它扩充成 V 的一个基$\alpha_1,\alpha_2,\cdots,\alpha_n$,则 $\mathcal{A}$ 在此基下的矩阵 $A=\begin{pmatrix}A_1 & A_3\\ 0 & A_2\end{pmatrix}$,其中 A_1 是 $\mathcal{A}|V_{\lambda_0}$在 V_{λ_0} 的一个基 $\alpha_1,\cdots,\alpha_r$ 下的矩阵,从而 $A_1=\lambda_0 I_r$.计算$|\lambda I-A|$.

*8. 提示:由于 V 是复数域上的 n维线性空间,因此 $\mathcal{A}$ 的特征多项式 $f(\lambda)=|\lambda I-A|=(\lambda-\lambda_1)^{l_1}\cdots(\lambda-\lambda_s)^{l_s}$,其中 $\lambda_1,\cdots,\lambda_s$ 是 $\mathcal{A}$ 的所有不同的特征值.显然,$l_1+\cdots+l_s=n$.设 $\dim V_{\lambda_i}=r_i$, $i=1,2,\cdots,s$.利用第7题的结论得,$r_i\leq l_i$, $i=1,2,\cdots,s$.然后利用 $\mathcal{A}$ 可对角化的第三个充分必要条件(关于特征子空间的维数).

9. 提示: $\mathcal{A}$ 是幂等变换 $\mathcal{A}^2 = \mathcal{A}$ $\mathcal{A}(\mathcal{A} - \mathcal{I}) = 0$ $\mathrm{Ker}[\mathcal{A}(\mathcal{A} - \mathcal{I})] = \mathrm{Ker}\,0 = V$. 注意 $(x, x - 1) = 1$, 利用本节定理 7 的结论.

*10. 提示: 设 V 是域 F 上 n 维线性空间, V 中取一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$. 构造 V 上的一个线性变换 $\mathcal{A}$, 使得 $\mathcal{A}(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)A$, 则 $A^2 = I$ $\mathcal{A}^2 = \mathcal{I}$. 然后类似于第 9 题的证法.

11. $f(x) = x - k$.

12. 提示: 由零化多项式的定义立即得出结论.

13. 提示: $A^2 = 0$.

14. 提示: $A^n = 0$.

*15. 提示: 先证 $\bar{\mathcal{A}}$ 是 V W 到 V W 的一个映射 (去说明对应法则 $\bar{\mathcal{A}}$ 与陪集 $\alpha + W$ 的代表的选择无关), 然后去证 $\bar{\mathcal{A}}$ 保持加法和纯量乘法两种运算.

习 题 9.6

1. 提示: 利用 §5 的推论 8 和 Hamilton - Cayley 定理.

2. 提示: 利用 Hamilton - Cayley 定理, 注意 A 的特征多项式的常数项等于 $(-1)^n|A|$.

3. (1) $\lambda^2 - 5\lambda + 6$; (2) $\lambda^2 - 2\lambda + 2$.

*4. 提示: 必要性. 用反证法. 假设 A 与 B 有公共特征值 λ_0, 则存在 $\alpha \in C^n$, $\beta \in C^m$ 且 $\alpha \neq 0$, $\beta \neq 0$, 使得 $A\alpha = \lambda_0\alpha$, $B\beta = \lambda_0\beta$. 令 $C = \alpha\beta'$, 去证 C 是 $AX - XB = 0$ 的非零解. 充分性. 假设 A 与 B 没有公共特征值, 则 A 的特征多项式 $f(x)$ 与 B 的特征多项式 $g(\lambda)$ 没有公共复根, 从而 $f(\lambda)$ 与 $g(\lambda)$ 互素, 于是存在 $u(\lambda), v(\lambda) \in C[\lambda]$, 使得

$$u(\lambda)f(\lambda) + v(\lambda)g(\lambda) = 1.$$

λ 用 A 代入, 利用 Hamilton - Cayley 定理, 可推出 $g(A)$ 可逆. 设 C 是 $AX - XB = 0$ 的一个解. 则 $AC = CB$. 去证 $g(A)C = Cg(B)$.

习 题 9.7

1. (1) $\lambda^2 - 1$; (2) $(\lambda - 2)^3$.

2. (1) $(\lambda - 1)^2$; (2) $(\lambda - 1)^3$;
   (3) $(\lambda - 3)^2(\lambda - 5)$; (4) $(\lambda - 3)^2(\lambda - 5)$.

3. 不相似.

4. 第 1 题中 (1) 小题的矩阵可对角化; (2) 小题中的矩阵不可对角化.

5. 第 2 题中各个 Jordan 形矩阵都不可对角化.

6. 提示: $A^m = I$, 由此得出 $\lambda^m - 1$ 是 A 的一个零化多项式. 考虑 $\lambda^m - 1$ 在复数域上的因式分解, 进而考虑 A 的最小多项式 $m(\lambda)$ 的因式分解.

7. 提示: 先求 A 的一个零化多项式, 进而考虑 A 的最小多项式 $m(\lambda)$. 由此判断 A 可对角化.

8. 提示: 方法同第 7 题, A 可对角化.

9. 提示: 先说明 A 的最小多项式 $m(\lambda)$ 与 $g(\lambda)$ 相伴, 从而 $m(\lambda)$ 在 Q 上不可约. 由于 $m(\lambda)$ 的次数 $r > 1$, 因此 A 不可对角化. 如果把 A 看成复矩阵, 去证 $m(\lambda)$ 可分解成 C 上不同的一

次因式的乘积,从而复矩阵 A 可对角化.

10. (1) 提示:先求 A 的特征多项式 $f(\lambda)$,利用 A 的最小多项式 $m(\lambda)$ 与特征多项式 $f(\lambda)$ 在 K 中有相同的根(重数可以不同),分析 $m(\lambda)$ 有几种可能情形,排除某些情形,最后可得 $m(\lambda)=(\lambda-1)^2(\lambda-2)$;

(2) 提示: $V=\mathrm{Ker}(A-I)^2 ī\ \mathrm{Ker}(A-2I)$.

$$\alpha\in \mathrm{Ker}(A-I)^2 \qquad (A-I)^2\alpha=0 \qquad (A-I)^2X=0,$$

其中 X 是 α 在基 $\alpha_1,\alpha_2,\alpha_3$ 下的坐标.由此可推出, $\alpha_1,\alpha_2-\alpha_3$ 是 $\mathrm{Ker}(A-I)^2$ 的一个基.类似地可求出 α_2 是 $\mathrm{Ker}(A-2I)$ 的一个基.

11. 提示:由已知条件得,

$$A\alpha_1=\alpha_2,\ A\alpha_2=\alpha_3,\cdots,\ A\alpha_{n-1}=\alpha_n,$$

$$A\alpha_n=-a_0\alpha_1-a_1\alpha_2-\cdots-a_{n-1}\alpha_n.$$

从而

$$A^2\alpha_1=A\alpha_2=\alpha_3,$$

$$A^3\alpha_1=A\alpha_3=\alpha_4,$$

$$\cdots\cdots$$

$$A^{n-1}\alpha_1=\alpha_n,$$

$$A^n\alpha_1=A\alpha_n,$$

$$A\alpha_n+a_{n-1}\alpha_n+\cdots+a_1\alpha_2+a_0\alpha_1=0.$$

即

$$A^n\alpha_1+a_{n-1}A^{n-1}\alpha_1+\cdots+a_1A\alpha_1+a_0\alpha_1=0.$$

令

$$g(\lambda)=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0,$$

则由上式得

$$g(A)\alpha_1=0.$$

去证 $g(A)\alpha_2=0,\cdots,g(A)\alpha_n=0$.从而 $g(A)=0$.因此 $g(\lambda)$ 是 A 的一个零化多项式.于是 $m(\lambda)\mid g(\lambda)$.假如 $\deg m(\lambda)<n$,去推出矛盾.因此

$$\deg m(\lambda)=n.$$

由此推出 $m(\lambda)=g(\lambda)$.即

$$m(\lambda)=\lambda^n+a_{n-1}\lambda^{n-1}+\cdots+a_1\lambda+a_0.$$

由于 $m(\lambda)\mid f(\lambda)$,且 $\deg m(\lambda)=n=\deg f(\lambda)$.从而可推出 $f(\lambda)=m(\lambda)$.

*12. 提示:只要证 $F[A]$ 的每一个非零元 $g(A)$ 可逆.由于 $g(A)\neq 0$,因此 $m(\lambda)8\ g(\lambda)$.利用不可约多项式的性质,可证出 $g(A)$ 可逆.

*13. 提示:设 A 的全部不同的特征值为 $\lambda_1,\cdots,\lambda_s$,其中 λ_i 的重数为 n_i, $i=1,\cdots,s$.于是存在域 F 上一个 n 级可逆矩阵 P,使得 $P^{-1}AP=\mathrm{diag}\{\lambda_1 I_{n_1},\cdots,\lambda_s I_{n_s}\}$,记作 D.令 $G=P^{-1}BP$.由 $AB=BA$ 可推出, $GD=DG$.据上册第 4 章习题 4.5 的第 18 题的结果, G 必为分块对角矩阵.设 $G=\mathrm{diag}\{G_1,G_2,\cdots,G_s\}$.考虑 G 与 $G_1,\cdots,G_s$ 的最小多项式之间的关系,可证 G_i 可对角化, $i=1,\cdots,s$.

习　题　9.8

1. (1) 幂零指数为 2;

(2) 提示:求出 $\mathrm{rank}\,B = 2$,因此 $BX = 0$ 的解空间的维数等于 $4 - 2 = 2$.从而 B 的 Jordan 标准形中 Jordan 块的总数为 2.由此和幂零指数可决定 B 的 Jordan 标准形.

2. 提示: B^k 也是幂零矩阵,从而 B^k 的特征多项式 $f(\lambda) = \lambda^n$.然后利用上册第 5 章§5 的命题 2 的结论.

3. 提示:必要性易证.充分性.由于 B 的特征值全为 0,因此 B 的最小多项式 $m(\lambda) = \lambda^l$,对某个正整数 l.

习 题 9.9

1. (1) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$; (2) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$;

(3) $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$; (4) $\begin{pmatrix} 3 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$;

(5) $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}$; (6) $\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}$.

2. (1) $(\lambda - 2)^2$; (2) $(\lambda - 1)\lambda^2$;
(3) $(\lambda - 1)^3$; (4) $(\lambda - 3)(\lambda + 1)^2$;
(5) $(\lambda - 1)^2(\lambda + 1)^2$; (6) $(\lambda - 2)^2$.

3. 提示:利用下述矩阵:

$$\begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \cdots & \cdots & & \cdots & \cdots \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

4. 提示:利用 A 的 Jordan 标准形以及第 3 题结论.

5. 提示:利用 A 的 Jordan 标准形.

6. 提示:利用 A 的 Jordan 标准形.

7. (1) 提示:先求出 A 的特征多项式 $f(\lambda) = \lambda^3 - \lambda^2 - \lambda + 1$.从而有 $A^3 = A^2 + A - I$.然后用数学归纳法证明 $A^k = A^{k-2} + A^2 - I$.

(2) 提示:利用第(1)小题的公式可以简便地计算出 A^{100}:

$$A^{100} = \begin{pmatrix} 1 & 0 & 0 \\ 50 & 1 & 0 \\ 50 & 0 & 1 \end{pmatrix}.$$

习 题 9.10

1. (1) 不是; (2) 是.

2. 提示:先求出 $f(\alpha_1), f(\alpha_2), f(\alpha_3)$,然后可求出

$$f(x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3) = 2x_1 - 3x_2 + x_3.$$

3. 提示:设 f 是满足题设条件的线性函数,求出 $f(\alpha_1), f(\alpha_2), f(\alpha_3)$,从而可求出 f 的表达式为

$$f(x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3) = -x_1 + 5x_2 + 4x_3.$$

*4. 提示:由于 $f_i \neq 0$,因此 $\mathrm{Ker}\, f_i \quad V, i = 1, 2, \cdots, s$.从而可利用习题 8.2 的第 10 题的结论.

5. 提示:设$(\beta_1, \beta_2, \beta_3) = (\alpha_1, \alpha_2, \alpha_3)A$,由已知条件可写出 A,证明 A 可逆,且求 A^{-1}.然后用定理 1 的结论得,$\beta_1, \beta_2, \beta_3$ 的对偶基 g_1, g_2, g_3 满足

$$(g_1, g_2, g_3) = (f_1, f_2, f_3)(A^{-1})',$$

其中

$$(A^{-1})' = \frac{1}{9}\begin{pmatrix} 2 & 1 & -2 \\ 1 & 2 & 2 \\ 2 & -2 & 1 \end{pmatrix}.$$

6. 提示:由于 $V = R[x]_3$,因此 $\dim V^* = \dim V = 3$.去证 f_1, f_2, f_3 线性无关,则 f_1, f_2, f_3 是 V^* 的一个基,显然 $1, x, x^2$ 是 V 的一个基,设它的对偶基为 $\tilde{f}_1, \tilde{f}_2, \tilde{f}_3$,据本节(9)式表明的结论得,$f_i = f_i(1)\tilde{f}_1 + f_i(x)\tilde{f}_2 + f_i(x^2)\tilde{f}_3, i = 1, 2, 3$.从而可求出

$$(f_1, f_2, f_3) = (\tilde{f}_1, \tilde{f}_2, \tilde{f}_3)\begin{pmatrix} 1 & 2 & -1 \\ \frac{1}{2} & 2 & \frac{1}{2} \\ \frac{1}{3} & \frac{8}{3} & -\frac{1}{3} \end{pmatrix}.$$

由于 $g_1(x), g_2(x), g_3(x)$的对偶基是 f_1, f_2, f_3,因此据本节定理 1 得

$$(g_1(x), g_2(x), g_3(x)) = (1, x, x^2)\begin{pmatrix} 1 & -\frac{1}{6} & -\frac{1}{3} \\ 1 & 0 & 1 \\ -\frac{3}{2} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

由此得出,$g_1(x) = 1 + x - \frac{3}{2}x^2$, $g_2(x) = -\frac{1}{6} + \frac{1}{2}x^2$, $g_3(x) = -\frac{1}{3} + x - \frac{1}{2}x^2$.

7. (1) 提示:利用线性映射的乘积还是线性映射;

(2) 提示:去证 $\mathcal{A}^*$ 保持加法和纯量乘法;

(3) 提示: $\mathcal{A}^*(f_1, f_2, \cdots, f_n) = (f_1\mathcal{A}, f_2\mathcal{A}, \cdots, f_n\mathcal{A})$, 而 $f_i\mathcal{A} = \sum_{j=1}^{n}[(f_i\mathcal{A})(\alpha_j)]f_j = \sum_{j=1}^{n}[f_i(\mathcal{A}\alpha_j)]f_j$. 又已知

$$\mathcal{A}(\alpha_1, \alpha_2, \cdots, \alpha_n) = (\alpha_1, \alpha_2, \cdots, \alpha_n)A,$$

因此可求出 $f_i\mathcal{A} = \sum_{j=1}^{n} a_{ij}f_j$.由此推出

$$A^{*}(f_1, f_2, \cdots, f_n) = (f_1, f_2, \cdots, f_n) A'.$$

8. (1) 提示: $W = \bigcap_{i=1}^{s} \mathrm{Ker} f_i$;

*(2) 提示:任取 V 的一个子空间 U,在 U 中取一个基 $\alpha_1, \alpha_2, \cdots, \alpha_m$,扩充成 V 的一个基 $\alpha_1, \cdots, \alpha_m, \alpha_{m+1}, \cdots, \alpha_n$,它的对偶基为 $f_1, f_2, \cdots, f_n$.任取 $\alpha \in V$,据本节(8)式表明的结论得,$\alpha = \sum_{i=1}^{n} f_i(\alpha)\alpha_i$.于是

$$\alpha \in U \qquad f_j(\alpha) = 0, \quad j = m+1, \cdots, n.$$

由此可得结论.

第 10 章　具有度量的线性空间

习　题　10.1

1. (1) 提示:按照定义去验证;

(2) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$;

(3) 提示:度量矩阵是满秩的;

(4) 提示:度量矩阵是对称的;

(5) $\alpha = (1, 0, 0, 1)'$(答案不唯一).

2. 提示: V 中取一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$,设 f 在此基下的度量矩阵为 A,设 α, β 在此基下的坐标分别为 X, Y.去证 $\alpha \in \mathrm{rad}_L V$ 当且仅当 X 属于 $A'Z = 0$ 的解空间 W.由于 $\alpha \longmapsto X$ 是 V 到 F^n 的一个同构映射,且在映射下,$\mathrm{rad}_L V$ 映成 W.因此

$$\dim \mathrm{rad}_L V = \dim W = n - \mathrm{rank}\, A' = \dim V - \mathrm{rank}_m f.$$

同理可证 $\dim \mathrm{rad}_R V = \dim V - \mathrm{rank}_m f$.

3. (1) 提示:按照线性映射的定义去证;

(2) 提示:去证 $\alpha \in \mathrm{rad}_L V$ 当且仅当 $\alpha \in \mathrm{Ker} L_f$;

(3) 提示:利用第 2 题的结论;

(4) 提示:利用第(3)小题的结论去证 f 非退化　　L_f 是单射;然后用第 9 章§2 的推论 4.

4. 提示:去求 f 在基 $E_{11}, E_{12}, \cdots, E_{1n}, \cdots, E_{n1}, \cdots, E_{nn}$ 下的度量矩阵.

5. (1) 提示:在 V 中取一个基 $\alpha_1, \alpha_2, \cdots, \alpha_n$,使得 f 在此基下的表达式为 $f(\alpha, \beta) = x_1 y_1 + x_2 y_2 + \cdots + x_r y_r$,其中 $r = \mathrm{rank}_m f$;

(2) 提示:取 $\xi = \left(\frac{2}{2}, \frac{2}{2}i, 0, \cdots, 0\right)$,$\eta = \left(\frac{2}{2}, -\frac{2}{2}i, 0, \cdots, 0\right)$.

6. 提示:由定义易证得结论.

7. 提示:按照子空间的充要条件去证.

8. (1) 提示:去证 $\alpha \in W^{\perp}$ 当且仅当 α 的坐标是某个齐次线性方程组的解空间;

(2) 提示:利用第(1)小题结论.

习　题　10.2

1. 是.提示:按照内积的定义逐条验证.关于正定性,把(α,α)的表达式配方.

2. 不是.提示:$f(A, B)$不满足正定性.

3. 提示:利用命题 1.

4. $\arccos\left(-\frac{1}{3}\right)$.

5. $\frac{2}{2}, \frac{6}{2}x, \frac{3\quad 10}{4}x^2 - \frac{10}{4}$.

6. $\alpha_1, \frac{10}{10}\alpha_2, -\frac{15}{3}\alpha_1 + \frac{15}{15}\alpha_2 + \frac{15}{3}\alpha_3$.

7. 提示:证法一.验证$(\beta_i, \beta_j) = \delta_{ij}, 1 \leq i, j \leq 3$.证法二.用命题 8.

8. (1) $(\alpha_1, \alpha_1) = 6, (\alpha_1, \alpha_2) = (\alpha_2, \alpha_1) = -2, (\alpha_1, \alpha_3) = (\alpha_3, \alpha_1) = 1$,

$(\alpha_2, \alpha_2) = 3, (\alpha_2, \alpha_3) = (\alpha_3, \alpha_2) = -2, (\alpha_3, \alpha_3) = 3$.

(2) $\alpha_1, \frac{1}{3}\alpha_1 + \alpha_2, \frac{1}{14}\alpha_1 + \frac{5}{7}\alpha_2 + \alpha_3$.

9. 提示:先求出 V 的一个标准正交基:$\varepsilon_1, \frac{1}{2}\varepsilon_2$,其中 $\varepsilon_1 = (1,0)', \varepsilon_2 = (0,1)'$.然后求出 $\alpha = (x_1, x_2)'$在 V 的标准正交基 $\varepsilon_1, \frac{1}{2}\varepsilon_2$ 下的坐标为$(x_1, \ 2x_2)'$.把 α对应到它的坐标的映射σ就是 V 到 R^2 的一个同构映射.

习　题　10.3

1. $\dim U^{\perp} = 2$, $U^{\perp}$的一个正交基是

$$\beta_1 = (0, -2, 1, 0), \quad \beta_2 = \left(2, -\frac{3}{5}, -\frac{6}{5}, 1\right).$$

提示:$\gamma \in U^{\perp} \quad \begin{pmatrix}\alpha_1 \\ \alpha_2\end{pmatrix}\gamma' = 0 \quad \gamma'$是 $\begin{pmatrix}\alpha_1 \\ \alpha_2\end{pmatrix}X = 0$ 的解.

2. $n-1$.　　3. 提示:显然 $U \quad (U^{\perp})^{\perp}$.然后去证 $\dim U = \dim(U^{\perp})^{\perp}$.

4. 提示:取 $U^{\perp}$的一个基 $\eta_1, \eta_2, \cdots, \eta_m$.令 $A = (\eta_1, \eta_2, \cdots, \eta_m)$,则 U 是齐次线性方程组 $A'X = 0$ 的解空间.

5. 提示:由于 $\eta_1, \eta_2, \cdots, \eta_m$ 是 U 的一个标准正交基,因此 $\alpha_1 = \sum_{i=1}^{m}(\alpha_1, \eta_i)\eta_i$.再去证 $(\alpha_1, \eta_i) = (\alpha, \eta_i)$.

6. $\alpha_1 = \left(-\frac{23}{29}, -\frac{48}{29}, -\frac{26}{29}\right)$.提示:利用第 5 题的结论;先求 U 的一个标准正交基 η_1,

η_2.

7. 提示:利用 $\alpha_1 \in U$ 是 α 在 U 上的正交投影当且仅当 $\alpha - \alpha_1 \in U^{\perp}$.

8. 提示:令 $W = \langle \varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m \rangle$,则 $V = W \oplus W^{\perp}$.任取 $\alpha \in V$,有 $\alpha = \alpha_1 + \alpha_2$, $\alpha_1 \in W$, $\alpha_2 \in W^{\perp}$.求出 $|\alpha_1|^2$,然后用勾股定理.

9. 提示:由于 $R[x]_4 = W \oplus W^{\perp}$, $\dim W = 1$,因此 $\dim W^{\perp} = 3$.由于 W 的一个基是 1,因此 $f(x) \in W^{\perp}$ 当且仅当 $(f,1) = \int_0^1 f(x)dx = 0$.分别考虑 $f(x)$ 是 1 次、2 次、3 次多项式的情形,可求出 $W^{\perp}$ 的一个基为

$$x - \frac{1}{2}, \quad x^2 - \frac{1}{3}, \quad x^3 - \frac{1}{4}.$$

10. 提示:$\dim W = n$,因此 $\dim W^{\perp} = n^2 - n$.易验证,$E_{11}, E_{22}, \cdots, E_{nn}$ 是 W 的一个标准正交基,于是 $A \in W^{\perp}$ 当且仅当 $(A, E_{ii}) = 0$, $i = 1, 2, \cdots, n$.由此可推出

$$W^{\perp} = \{A = (a_{ij}) \in M_n(R) \mid a_{11} = a_{22} = \cdots = a_{nn} = 0\}.$$

于是 $E_{ij} \in W^{\perp}$,其中 $i \neq j$, $i, j = 1, 2, \cdots, n$,去证它们组成 $W^{\perp}$ 的一个标准正交基.

习　题　10.4

1. 提示:设 ξ 是 A 的属于特征值 λ_1 的特征向量,用两种方法计算$(A\xi, A\xi)$.

2. 提示:注意利用 $P\alpha = (\alpha, \eta)\eta$(这是习题 10.3 第 5 题的结论),去证$(A\alpha, A\beta) = (\alpha, \beta)$.再去证 A 是满射;任取 $\gamma \in V$,令 $\alpha = \gamma - 2P\gamma$,则 $A\alpha = \gamma$.

为了证明 A 是第二类的,在 $\langle \eta \rangle^{\perp}$ 中取一个基 $\alpha_2, \cdots, \alpha_n$,去求 A 在 V 的一个基 $\eta, \alpha_2, \cdots, \alpha_n$ 下的矩阵,然后计算它的行列式.

3. 提示: $V_1^{\perp}$ 是 1 维的,设 $V_1^{\perp} = \langle \eta \rangle$. V_1 中取一个基 $\alpha_1, \alpha_2, \cdots, \alpha_{n-1}$,则 $\alpha_1, \alpha_2, \cdots, \alpha_{n-1}, \eta$ 是 V 的一个基.设 P 是 V 在 $\langle \eta \rangle$ 上的正交投影,则

$$A\alpha_i = \alpha_i = (I - 2P)\alpha_i, \quad i = 1, 2, \cdots, n-1.$$

由于 $$(A\eta, A\alpha_i) = (\eta, \alpha_i) = 0,$$

又 $$(A\eta, A\alpha_i) = (A\eta, \alpha_i),$$

因此 $$(A\eta, \alpha_i) = 0, i = 1, 2, \cdots, n-1.$$

从而 $A\eta \in V_1^{\perp} = \langle \eta \rangle$.由于 A 是正交变换,因此 $|A\eta| = |\eta|$,从而 $A\eta = \pm \eta$.由于 $\dim V_1 = n-1$,因此 $A\eta \neq \eta$.从而 $A\eta = -\eta = (I - 2P)\eta$,因此 $A = I - 2P$,从而 A 是镜面反射.

4. 提示:必要性显然,充分性,去证 A 保持内积:用两种方法计算 $|A(\alpha + \beta)|^2$,即可证得 $(A\alpha, A\beta) = (\alpha, \beta)$.

5. 提示:正交变换 A 在 V 的任意一个标准正交基下的矩阵 A 是正交矩阵.设 $A = (a_{ij})$.由正交矩阵的列向量组的特点,可得出点 $P(a_{11}, a_{21})$和点 $M(a_{12}, a_{22})$都在平面直角坐标系的单位圆上,且 $\overline{OP} \perp \overline{OM}$.设点 P, M 分别在角 θ, φ 的终边上,则 $\varphi = \theta + \frac{\pi}{2}$ 或 $\theta = \varphi + \frac{\pi}{2}$.由此可得出 A 有且只有两种可能.

6. 提示:令 $\eta = k(\alpha - \beta)$且 $|\eta| = 1$,求出 k.然后设 A 是关于超平面 $\langle \eta \rangle^{\perp}$ 的镜面反射,即 $A = I - 2P$,其中 P 是 V 在 $\langle \eta \rangle$ 上的正交投影,去证 $A\alpha = \beta$.

7. 提示:设 λ_0 是正交矩阵 A 的特征多项式的任一复根,把 A 看成复矩阵,则 λ_0 是 A 的特征值,从而有特征向量 α,使得 $A\alpha=\lambda_0\alpha$.取共轭转置得,$\bar{\alpha}'A'=\bar{\lambda}_0\bar{\alpha}'$.由这两个式子得,$\bar{\alpha}'A'A\alpha=\bar{\lambda}_0\lambda_0\bar{\alpha}'\alpha$.由此推出 $|\lambda_0|=1$.

8. 提示:充分性.由命题 5 立即得到,必要性.对欧几里得空间的维数 n 用数学归纳法.在讨论 n 维欧几里得空间 V 上的对称变换 A 时,取 A 的一个特征值 λ_1,设 η_1 是 A 的属于 λ_1 的一个特征向量,且 $|\eta_1|=1$.则 $V=\langle\eta_1\rangle\oplus\langle\eta_1\rangle^{\perp}$.显然 $\langle\eta_1\rangle$ 是 A 的不变子空间,据命题 6 得,$\langle\eta_1\rangle^{\perp}$ 也是 A 的不变子空间,于是 $A|\langle\eta_1\rangle^{\perp}$ 是 $\langle\eta_1\rangle^{\perp}$ 上的线性变换且是对称的.对 $\langle\eta_1\rangle^{\perp}$ 用归纳假设,即可证得结论.

习　题　10.5

1. $|\alpha|=3$, $|\beta|=2$, α 与 β 的夹角 $\langle\alpha,\beta\rangle=\frac{\pi}{6}$.

2. $\eta_1=\left(\frac{2}{2},-\frac{2}{2}\right)$, $\eta_2=\left(\frac{1+i}{2},\frac{1+i}{2}\right)$.

3. $\beta_1=(1,-1,1)$, $\beta_2=\left(\frac{2-i}{3},\frac{1+i}{3},\frac{-1+2i}{3}\right)$.

4. $(e^{i\theta})$,其中 θ 是实数.　　5. 提示:对于复矩阵 A,有 $|\overline{A}|=\overline{|A|}$.

6. 提示:设 ξ 是 A 的属于特征值 λ_1 的特征向量.用两种方法计算 $(A\xi, A\xi)$.

7. $\begin{pmatrix} a & b+ci \\ b-ci & d \end{pmatrix}$,其中 a, b, c, d 是任意实数.

*8. 提示:设 $\eta_1,\eta_2,\cdots,\eta_n$ 是 V 的一个标准正交基,

$$A(\eta_1,\eta_2,\cdots,\eta_n)=(\eta_1,\eta_2,\cdots,\eta_n)A,$$

则 $A\eta_i$ 在基 $\eta_1,\eta_2,\cdots,\eta_n$ 下的坐标的第 j 个分量 $a_{ji}=(A\eta_i,\eta_j)$.

*9. 提示:设 ξ 是 A 的属于特征值 λ_1 的特征向量,用两种方法计算 $(A\xi,\xi)$.

*习　题　10.6

1. (1) 提示:只要证 f 是非退化的对称双线性函数;

(2) 提示:按标准正交基的定义验证;

(3) 提示:$A=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$;

(4) 提示:设 $\alpha=(\varepsilon_1,\varepsilon_2)X$, $\beta=(\varepsilon_1,\varepsilon_2)Y$,则 $T\alpha$, $T\beta$ 在基 $\varepsilon_1,\varepsilon_2$ 下的坐标分别为 TX, TY.从而

$$f(T\alpha,T\beta)=(TX)'A(TY)=X'T'ATY$$

由于 $f(\alpha,\beta)=X'AY$.因此 T 是正交变换当且仅当

$$X'T'ATY=X'AY,\quad \forall\ x,y\in R^2.$$

而后者等价于 $T'AT=A$.从而只要去证 $T'AT=A$ 就可得出 T 是正交变换;

(5) 属于 2 + 1 的特征向量是 $k\begin{pmatrix}1\\1\end{pmatrix}$, k 是任意非零实数;属于 2 - 1 的特征向量是

$l\begin{pmatrix}1\\-1\end{pmatrix}$, l 是任意非零实数.

2. 提示:设 $\beta=\sum_{i=1}^{n}b_i\alpha_i$,去计算 $f(\beta,\alpha_j)$.

3. 提示: f 在辛基下的矩阵 A 为

$$\begin{pmatrix}0 & I_r\\-I_r & 0\end{pmatrix}.$$

设 α,β 在辛基下的坐标分别为 X,Y,则 $f(\alpha,\beta)=X'AY$,

$$f(B\alpha,B\beta)=(BX)'A(BY)=X'B'ABY.$$

线性变换 B 是辛变换当且仅当 $f(B\alpha,B\beta)=f(\alpha,\beta)$.由此可推出 $B'AB=A$.

参 考 文 献

1. 北京大学数学系几何与代数教研室代数小组.高等代数.第二版.北京:高等教育出版社,1988
2. Kostrikin A I.代数学引论(上册).张顺燕,蓝以中,丘维声译.北京:高等教育出版社,1988
3. Kostrikin A I & Manin Yu I. Linear Algebra And Geometry. Translated by Alferieff M E. New York. Gordon And Breach Science Publishers,1989
4. Hoffman K & Kunze R. Linear Algebra. Second Edition. New Jersey. Prentice-Hall Inc,1971
5. Postnikov M.几何讲义,第二学期,线性代数和微分几何.陈维桓,石生明译.北京:高等教育出版社,1992
6. Postnikov M.几何讲义,第一学期,解析几何.周友成译.北京:高等教育出版社,1992
7. 许以超.代数学引论.上海:上海科技出版社,1965
8. 张禾瑞,郝炳新.高等代数.第三版.北京:高等教育出版社,1984
9. 李炯生,查建国.线性代数.合肥:中国科学技术大学出版社,1989
10. H B普罗斯库烈柯夫.线性代数习题集.周晓钟译.北京:人民教育出版社,1981
11. Van der Waerden B L.代数学 I.丁石孙,曾肯成,郝炳新译.万哲先校.北京:科学出版社,1963
12. Jacobson N. Basic Algebra I. San Francisco. W H Freeman and Company,1974
13. 聂灵沼,丁石孙.代数学引论.北京:高等教育出版社,1988
14. 莫宗坚,蓝以中,赵春来.代数学(上册).北京:北京大学出版社,1986
15. 张恭庆,林源渠.泛函分析讲义(上册).北京:北京大学出版社,1987
16. 丘维声.解析几何.北京:北京大学出版社,1988